# Distinct Roles of Dopamine D1 and D2 Receptor-expressing Neurons in the Nucleus Accumbens for a Strategy Dependent Decision Making

**Tadaaki Nishioka[1,*], Tom Macpherson[1], Kosuke Hamaguchi[2] and Takatoshi Hikida[1,3,*]**

[1]Laboratory for Advanced Brain Functions, Institute for Protein Research, Osaka University;

[2]Department of Biological Sciences, Graduate School of Medicine, Kyoto University;

[3]Lead Contact

*Correspondence: nishioka@protein.osaka-u.ac.jp (T.N.), hikida@protein.osaka-u.ac.jp (T.H.)

## SUMMARY

To optimize decision making, animals need to execute not only a strategy to choose a good option but sometimes also one to avoid a bad option. A psychological study indicates that positive and negative information is processed in a different manner in the brain. The nucleus accumbens (NAc) contains two different types of neurons, dopamine D1 and D2 receptor-expressing neurons which are implicated in reward-based decision making and aversive learning. However, little is known about the neural mechanisms by which D1 or D2 receptor-expressing neurons in the NAc contribute to the execution of the strategy to choose a good option or one to avoid a bad option under decision making. Here, we have developed two novel visual discrimination tasks for mice to assess the strategy to choose a good option and one to avoid a bad option. By chemogenetically suppressing the subpopulation of the NAc neurons, we have shown that dopamine D2 receptor-expressing neurons in the NAc selectively contribute to the strategy to avoid a bad option under reward-based decision making. Furthermore, our optogenetic and calcium imaging experiments indicate that dopamine D2 receptor-expressing neurons are activated by error choices and the activation following an error plays an important role in optimizing the strategy in the next trial. Our findings suggest that the activation of D2 receptor-expressing neurons by error choices through learning enables animals to execute the appropriate strategy.

## INTRODUCTION

Even seemingly similar decisions could be mediated by different cognitive processes. For example, selecting a food for dinner could be mediated by either picking up one's favorite food or avoiding the one with a bad experience. Although selecting a better choice and avoiding a bad choice could lead to a similar choice behavior, a psychological study indicates that their underlying processes in the brain could be different (Tversky and Kahneman, 1981). A number of studies have proposed the basal ganglia circuitry is implicated in decision making (Cohen et al., 2012; Frank et al., 2004; Hamid et al., 2015; Hikida et al., 2010; Iino et al., 2020; Maia and Frank, 2011; Samejima et al., 2005; Schultz et al., 1997; Stephenson-Jones et al., 2016; Tai et al., 2012; Yagishita et al., 2014; Zalocusky et al., 2016). Dopamine has been known as an important modulator of basal ganglia functions (Ferenczi et al., 2016; Gerfen and Surmeier, 2011; Parker et al., 2018; Shen et al., 2008; Surmeier et al., 2009). The nucleus accumbens (NAc), one of the major projection targets of dopamine neurons, is known to play an important role in reward-based decision making and aversive learning (Berridge and Kringelbach, 2008; Carlezon and Thomas, 2009; Hikida et al., 2010; Iino et al., 2020; Lee, 2013; Yawata et al., 2012; Zalocusky et al., 2016). The medium spiny projection neurons (MSN) in the NAc are divided into two major subpopulations, one that expresses the dopamine D1 receptor (D1-MSN) and the other that expresses the dopamine D2 receptor (D2-MSN). Because D1 and D2 receptors have antagonistic response to dopamine, D1- and D2-MSN were thought to have different contribution to the choice behavior. Previous studies showed that D1-MSN and D2-MSN in the NAc play distinct roles in a context dependent manner (Aragona et al., 2006; Hikida et al., 2010; Lobo et al., 2010; Macpherson and Hikida, 2018; Yawata et al., 2012), however, it is not well understood how and whether different strategies, such as selecting a better choice or avoiding a bad choice, are differentially represented in these two distinct classes of neurons in NAc.

In this study, we have developed a pair of visual discrimination learning tasks to assess the strategy dependent decision making in mice. While in the visual discrimination-based cue-guided attendance learning (VD-Attend), mice need to acquire the strategy to choose a good option. In contrast, in the visual discrimination-based cue-guided avoidance learning (VD-Avoid), mice need to acquire the strategy to avoid a bad option. To establish the contribution of these two subpopulations in these different tasks, we chemogenetically manipulated the neuronal activities of D1-MSN or D2-MSN in the NAc while mice

55 performing the VD-Attend or VD-Avoid tasks. Moreover, we performed optogenetic manipulation to test
56 whether the NAc contributes to the execution of the strategy in a timing-specific manner. We further
57 performed in vivo calcium imaging with a miniature microscope to investigate the underlying neural
58 representations of D1-MSN and D2-MSN. Our results show that the majority of D1-MSN in the NAc is
59 recruited by correct choice and the activity of D1-MSN is important for the motivational control and keeping
60 the same strategy. By contrast, D2-MSN in the NAc plays a selective role in executing the strategy to avoid a
61 bad option in VD-Avoid task. Furthermore, we found that the vast majority of D2-MSN are activated by error
62 choices and the activation of D2-MSN after error is necessary for optimizing the strategy in next trial. Our
63 findings suggest that the activation of D2-MSN by error choices through learning enables animals to execute
64 the appropriate strategy.
65
66
67 **RESULTS**
68 **Mice Can Acquire Both the Strategy to Choose a Good Option and One to Avoid a Bad Option**
69 To study the neural mechanisms underlying the strategy to choose a good option and one to avoid a bad option,
70 we designed two novel visual discrimination learning tasks, VD-Attend and VD-Avoid task on a touchscreen-
71 based operant chamber (**Figure 1A**). In these tasks, two visual cues were presented on a touchscreen after trial
72 initiation (**Figure 1B**). If mice responded to a correct cue, the reward was delivered in the magazine. If mice
73 responded to an incorrect cue, they were punished with a 5-sec time-out. In the VD-Attend task, while the
74 visual cue A (e.g. marble) was presented as a correct cue in every trial, a random image was presented as an
75 incorrect cue. Therefore, mice were required to execute a strategy to choose cue A in order to obtain a reward
76 (**Figure 1A**). By contrast, in the VD-Avoid task, the visual cue B (e.g. fan) was presented as an incorrect cue
77 and a random image was presented as a correct cue. Therefore, mice cannot directly learn the correct cue, but
78 instead, mice could obtain a reward by avoiding the cue B (**Figure 1A**). Next, we trained C57BL/6J mice to
79 the VD-Attend or VD-Avoid task and confirmed that mice can perform both tasks well within a few weeks of
80 training (**Figure 1C**). The number of sessions and total errors to criterion in the VD-Avoid task were larger
81 than that in the VD-Attend task (**Figure 1D and 1E**), suggesting that acquisition of the strategy to avoid a bad

82 option is more difficult than that to choose a good option. However, the performances measured by the
83 percentage of correct choice were comparable in the VD-Attend and VD-Avoid tasks across sessions (**Figure**
84 **1F and 1G**). Correct latency and reward latency were also comparable in the VD-Attend and VD-Avoid
85 (**Figure 1H and 1I**). A number of studies have reported the phenomenon of post-error slowing in which the
86 subjects tend to slow down the response after an error trial (Houtman et al., 2012; Notebaert et al., 2009;
87 Rabbitt and Rodgers, 1977; Ruitenberg et al., 2014). We also tested whether the recent history on the previous
88 trial affected the response latency on the current trial. We found that mice tended to respond slowly on the next
89 trial after making an error response in the VD-Avoid task (**Figure S1**), suggesting that response latencies were
90 modulated by recent history. Together, these data show that mice can acquire both strategies of attending and
91 avoiding specific visual cues in the touchscreen-based operant tasks, which provide a framework for studying
92 the neural mechanisms underlying two different forms of strategies.
93

94 **Chemogenetic Suppression of D2-MSN in the NAc Selectively Impair the Ability to Execute the Strategy**
95 **to Avoid a Bad Option**

96 Given that dopaminergic medication impaired the ability to avoid a bad option option (Frank et al., 2004,
97 2007), and the NAc, one of the major targets of the dopamine neurons, plays an important role in decision
98 making (Berridge and Kringelbach, 2008; Carlezon and Thomas, 2009; Hikida et al., 2010; Iino et al., 2020;
99 Lee, 2013; Yawata et al., 2012; Zalocusky et al., 2016), we first investigated whether D1-MSN and D2-MSN
100 in the NAc contribute to the performance in VD-Attend and VD-Avoid tasks. We bilaterally injected an adeno-
101 associated virus (AAV) carrying inhibitory designer receptor exclusively activated by a designer drug
102 (iDREADD, hM4Di) fused to mCherry (Armbruster et al., 2007) in a Cre dependent manner into the NAc of
103 D1-Cre or D2-Cre mice (**Figure 2A**). The hM4Di can robustly suppress the neuronal activities via cell
104 signaling when an artificial ligand clozapine-N-oxide (CNO) is administered (Wulff and Arenkiel, 2012). To
105 test whether D1-MSN or D2-MSN in the NAc is involved in executing the strategy, we trained mice
106 expressing hM4Di-mCherry in D1-MSN or D2-MSN of the NAc to the VD-Attend or VD-Avoid task (**Figure**
107 **2A**). After mice were well-trained, we intraperitoneally administered CNO at the session 4 and 6 in probe test
108 (**Figure 2A**). In this study, we mainly targeted the core region of the NAc because recent studies indicate that

109 the dopamine neurons are heterogeneous and the dopamine neurons projecting to the core but not shell region
110 encode the reward prediction error information (Iino et al., 2020; Jong et al., 2019; Parker et al., 2016;
111 Roitman et al., 2008). We histologically confirmed that hM4Di-mCherry expression in the injection site were
112 mainly restricted to the NAc (**Figure 2B**). Moreover, in D1-Cre mice, hM4Di-mCherry-positive axon
113 terminals were observed in the ventral pallidum (VP) and substantia nigra pars reticulata (SNr), while in D2-
114 Cre mice, hM4Di-mCherry-positive axon terminals were observed only in the VP but not SNr (**Figure S2**),
115 which are consistent with the canonical targets of D1-MSN or D2-MSN in the NAc (Lu et al., 1997; Zhou et
116 al., 2003). These data demonstrates that the viral transduction allowed us to selectively express hM4Di-
117 mCherry in D1-MSN or D2-MSN in the NAc. A previous study also shows that CNO can functionally
118 suppress the neuronal activity of the NAc (Luo et al., 2018). Next, we examined how chemogenetic
119 suppression of D1-MSN or D2-MSN in the NAc via CNO influenced the performance of the VD-Attend or
120 VD-Avoid task across sessions (**Figure 2C**). Chemogenetic suppression of D1-MSN in the NAc reversibly
121 decreased the performance of both the VD-Attend and VD-Avoid task (**Figure 2D**). Interestingly,
122 chemogenetic suppression of D2-MSN in the NAc selectively decreased the performance of the VD-Avoid but
123 not VD-Attend task (**Figure 2D**), suggesting that the neuronal activities of D2-MSN in the NAc is necessary
124 for avoiding a bad option but not choosing a good option under decision. The extent of decreased performance
125 in D1-Cre mice was comparable to that in D2-Cre mice in the VD-Avoid task (**Figure 2E**). Moreover, we
126 confirmed that CNO itself did not affect the performance of the VD-Attend or VD-Avoid task though CNO
127 treatment slightly improved the performance of the VD-Avoid task in control group (**Figure 2D**). Considering
128 that the NAc plays an important role in motivational control (Aberman et al., 1998; Berridge, 2007; Delgado,
129 2007; Gallo et al., 2018; Roitman et al., 2005; Salamone et al., 2007; Schultz et al., 1992) and that a decrease
130 in the performance in the reward-based decision making could be associated with motivation deficits, we
131 analyzed the effect of chemogenetic manipulation on motivation. We found that chemogenetic suppression of
132 D1-MSN in the NAc decreased the number of earned rewards and increased reward latency, both of which
133 indicated decreased motivation (**Figure 2F and 2G**). On the other hand, chemogenetic suppression of D2-
134 MSN in the NAc did not affect the motivation indices (**Figure 2F and 2G**), suggesting that a decrease in
135 performance of the VD-Avoid task are attributed to the impaired ability to avoid a bad option. Furthermore, we

136 analyzed whether the outcome of the previous trial affected the performance of the current trial. The
137 chemogenetic suppression of D1-MSN in the NAc resulted in a decrease of performance following a correct
138 trial (**Figure S3A and S3B**), suggesting that D1-MSN in the NAc plays an important role in consecutively
139 making a correct choice. On the other hand, we could not detect the effect of the recent history on the
140 behavioral performance in D2-Cre or control group (**Figure S3C-S3F**). Collectively, our data indicate that D1-
141 MSN and D2-MSN in the NAc contribute to the choice behavior differently; D1-MSN contributes to both the
142 VD-Attend and VD-Avoid tasks, and D2-MSN selectively involved in the VD-avoid task.
143

144 **Optogenetic Inhibition of D2-MSN in the NAc during Outcome Period on the Error Trial Is Sufficient to**
145 **Impair the Ability to Execute the Strategy to Avoid a Bad Option on the Next Trial**

146 We have shown that D2-MSN in the NAc causally contributes to the strategy to avoid a bad option (**Figure 2**).
147 To identify a critical time window in which D2-MSN exerts influence on the choice behavior, we
148 optogenetically inhibited the neuronal activity of D2-MSN in the NAc in a time-locked manner. We first
149 bilaterally expressed the light-driven outward proton pump, archaerhodopsin (ArchT) fused to tdTomato (Han
150 et al., 2011) in D2-MSN in the NAc and implanted optic fibers above the NAc (**Figure 3A, 3B, and S4**). The
151 optical stimulation can suppress the activities of ArchT expressing neurons (Kang et al., 2017). We separated
152 the task into 3 time periods (ITI period; the last 5 sec of an inter-trial interval (ITI) to trial initiation. Cue
153 period; the time from trial initiation to the response. Outcome period; 5 sec after the response, **Figure 3C**) to
154 tested whether D2-MSN activity in different time periods have different effects on the performance of the VD-
155 Avoid task. We found that optogenetic inhibition of D2-MSN in the NAc during Outcome period selectively
156 decreased the performance but not in ITI or Cue period stimulation trials (**Figure 3D-3F**). Furthermore, a
157 decrease in performance by optogenetic inhibition during Outcome period was observed on the next trial after
158 making an error choice on the previous trial but not after making a correct choice (**Figure 3G and 3H**). The
159 recent history-dependent decline in the performance was observed only when optogenetic inhibition was
160 applied to Outcome period but not ITI or Cue period (**Figure S5A-S5D**). These results suggest that the
161 neuronal activity of D2-MSN in the NAc immediately after making an error response plays an important role
162 in avoiding a bad option on the next trial. In addition, since response latencies were modulated by recent

163 history (**Figure S1**) and optogenetic inhibition of D2-MSN in the NAc affect the performance in the VD-
164 Avoid task in a history dependent manner (**Figure 3G**), we analyzed whether optogenetic inhibition of D2-
165 MSN in the NAc affected response latencies in the VD-Avoid task. Interestingly, optogenetic inhibition of D2-
166 MSN in the NAc sped up the response on the current trial after having committed an error on the previous trial
167 (**Figure S6A-S6F**), suggesting that the neuronal activation of D2-MSN in the NAc immediately after making
168 an error response contributes to post-error slowing. Taken together, optogenetic inhibition of D2-MSN in the
169 NAc during Outcome period in the error trial selectively decreases the behavioral performance on the next
170 trial.
171

172 **D2-MSN in the NAc Are Dominantly Activated after Making an Error Response than D1-MSN**

173 We have shown that inhibition of D2-MSN in the NAc during Outcome period was sufficient to decrease the
174 behavioral performance in the VD-Avoid task (**Figure 3**). Next, we examined whether D2-MSNs are activated
175 by error response and how the neuronal activity differs between D1-MSN and D2-MSN. To investigate the
176 activity of D1-MSN or D2-MSN in the NAc, we performed calcium imaging of D1-MSN or D2-MSN in the
177 NAc at a single-cell resolution using a miniature microscope. We injected AAV carrying a fluorescent calcium
178 indicator, jGCaMP7f (Dana et al., 2019) in a Cre dependent manner into the NAc of D1-Cre or D2-Cre mice
179 and implanted a gradient-index (GRIN) lens above the NAc (**Figure 4A, 4B, and S7**). Using a constrained
180 non-negative matrix factorization method to the microendoscopic images (CNMFe; Zhou et al., 2018), we
181 extracted the calcium activities from individual D1-MSNs or D2-MSNs from freely moving mice performing
182 the VD-Avoid task (**Figure 4C and 4D; Movies S1 and S2**). We identified 266 cells from D1-Cre mice and
183 194 cells from D2-Cre mice, respectively. Some of cells were reliably activated after making an error response
184 but not a correct response (**Figure 4E**). Others showed a transient activation following a correct response
185 (**Figure 4F**). We have shown that D1-MSN and D2-MSN contribute to the behavioral performance in a recent
186 history dependent manner (**Figure S3A, S3B, and 4G**). Therefore, we focused on the difference in the calcium
187 activity between correct and error trials and calculated the area under the curve receiver operating
188 characteristic (auROC) of z-scored activities in the error and correct trials. To classify the calcium activities of
189 D1-MSN and D2-MSN in the NAc, we performed the principal component analysis (**Figure 4G**). In D2-MSN,

60% of identified cells (117 out of 194 cells) were classified as Type I, which showed sustained activation after making an error choice while 40% of identified cells (77 out of 194 cells) were classified as Type II, which were activated in a correct trial (**Figure 4H**). In D1-MSN, 44% of identified cells (116 out of 266 cells) were classified as Type I while 56% of identified cells (150 out of 266 cells) were classified as Type II (**Figure 4H**). Moreover, to investigate when D1-MSN or D2-MSN in the NAc showed statistically significant calcium activity between correct and error trials, we defined cell with statistically significant activity as responsive cells and calculated their proportions (**Figure 4I and S8**). As a result, we found that the proportion of responsive cells statistically increased in Outcome period compared to ITI and Cue period (**Figure 4J**). This result was consistent with the previous results by optogenetics in which inhibition of D2-MSN during Outcome period selectively decreased the behavioral performance (**Figure 3D-3F**). Furthermore, we analyzed the calcium dynamics of cells with statistically significant activity during Outcome period (**Figure 4K and S9**) and found that the proportion of Type I (Error Type) in D2-MSN was more dominant than D1-MSN (**Figure 4L**). These results were also consistent with the results by optogenetics in which inhibition of D2-MSN in the error trials decreased the performance on the next trial (**Figure 3G and 3H**). Although Type I and Type II were intermingled in both D1-MSN and D2-MSN (**Figure S10A-S10D**), D2-MSNs of the same valence existed closer to each other than those of opposite valence, suggesting that D2-MSN but not D1-MSN formed spatial clustering of cells with similar functional properties (**Figure S10E and S10F**). Collectively, the majority of D2-MSNs in the NAc was activated by error choice in the VD-Avoid task, corroborating our hypothesis that neuronal activation of D2-MSN in the NAc by error detection plays an important role in avoiding a bad option. On the other hand, the majority of D1-MSNs in the NAc was activated by correct choice. Given that inhibition of D1-MSN in the NAc decreased the performance on the next trial after making a correct choice, it could be that neuronal activation of D1-MSN in the NAc contributes to keep same strategy with confidence.

## Neuronal Activation of D2-MSN in the NAc by Error Choice Were Acquired through Learning in the VD-Avoid Task

To investigate if error-correct selective activity during the Outcome period depends on the learning, we next examined whether neuronal activation of MSN in the NAc by error or correct choices in the VD-Avoid task depended on the learning. To investigate how the neuronal activity of D1-MSN or D2-MSN in the NAc changed through learning at a single-cell level, we tracked the same neurons before and after learning (**Figure 5A**) with a cell registration methods (Sheintuch et al., 2017). As the training progressed, D1-Cre and D2-Cre mice similarly improved their performances in the VD-Avoid task (**Figure S11**). In the criterion session, we imaged the activities of 460 cells (D1-Cre, 266 cells; D2-Cre, 194 cells) and identified 239 neuron pairs (D1-Cre, 103 pairs; D2-Cre, 136 pairs) from these mice. A subset of cells activated by error choices (Type I) in the criterion session (expert) had a weak response to error choices in the first session (novice) in the VD-Avoid task (**Figure 5B**). Similarly, we also found that some of the cells activated by correct choices (Type II) in expert mice had a weak response to correct choices in novice mice in the VD-Avoid task (**Figure 5C**). Next, to investigate the change in the calcium activity against error or correct choices through learning, we analyzed the difference in the calcium activities in novice and expert mice (**Figure 5D**). By tracking the same neurons through training, we found that the activity of D2-MSN in expert mice positively correlated with that in novice mice, whereas we could not detect any correlation in D1-MSN imaged in the VD-Avoid task (**Figure 5D**). Furthermore, we classified the type of neuron pairs and quantified the proportion of cells that become responsive Type I or Type II from non-responsive in D1-Cre or D2-Cre mice. The results showed that the majority (13.6%) of D1-MSN acquired the correct choice selectivity after learning (**Figure 5E and S12**). In contrast, the majority (16.8%) of D2-MSN acquired the error choice selectivity after learning (**Figure 5E and S12**). Moreover, we revealed that the proportion of cells that become responsive Type I from non-responsive were significantly higher in D2-MSN than D1-MSN (**Figure 5F**). In contrast, the proportion of cells that become responsive Type II from non-responsive were significantly higher in D1-MSN than D2-MSN (**Figure 5F**). These results indicate that the neuronal activation of D1-MSN and D2-MSN by outcome depends on the learning. The acquisition of responsiveness by learning of these neurons is likely to contribute to the high performance in the VD-Avoid task.

243 **DISCUSSION**

244 Although dopamine signaling plays an important role in strategy-dependent decision making (Frank et al.,
245 2004, 2007), it has been poorly understood how the neural substrates targeted by dopamine contribute to
246 reinforcement learning and execution of the strategy under the decision. Here we have established novel visual
247 discrimination tasks in mice to assess the neural mechanisms underlying the strategy-dependent decision
248 making. Our data show that neuronal activation of D2-MSN in the NAc by error choices is essential for
249 executing the strategy to avoid a bad option. Previous studies have shown that when the dopamine
250 concentration in a brain is increased by L-dopa treatment, the subjects cannot avoid a bad option but still can
251 choose a good choice (Frank et al., 2004, 2007). L-dopa treatment can decrease the excitability of D2-MSN
252 because the dopamine D2 receptor is coupled to Gi signaling (Cepeda et al., 1993; Stoof and Kebabian, 1984;
253 Surmeier et al., 2007). Given that our chemogenetic and optogenetic suppression of D2-MSN in the NAc
254 disrupted the strategy to avoid a bad option but not that to choose a good option, we concluded that
255 hypoactivity of D2-MSN in the NAc caused the subjects not to execute the strategy to avoid a bad option.
256 Although overactivation of D1-MSN in the striatum including the NAc could also contribute to disability to
257 avoid a bad option directly or indirectly (e.g. via local D2-MSN to D1-MSN synapse; Matamales et al., 2020),
258 our study demonstrates downstream neural mechanisms of dopamine signaling by which the neuronal activity
259 of D2-MSN in the NAc selectively contributes to avoid a bad option but not choose a good option.

260 Our calcium imaging data suggest that as individual D2-MSNs are reliably activated by error choices,
261 mice efficiently avoid a bad option. Although individual D1-MSNs lost the response selectivity through
262 learning (**Figure 5D**), given that chemogenetic suppression of D1-MSN decreased the performance in the VD-
263 Avoid task as well as the VD-Attend task (**Figure 2D**) and some D1-MSNs were recruited by correct choices
264 after learning (**Figure 5E and 5F**), we expect that the neuronal activity of D1-MSN also contributes to the
265 performance in the VD-Avoid task. Our calcium imaging data also showed that the activity of D2-MSN in
266 expert mice positively correlated with that in novice mice, whereas we could not detect any correlation in D1-
267 MSN imaged in the VD-Avoid task. Given that the VD-Avoid task required mice to avoid a specific cue (e.g.
268 flag) associated with negative outcome (e.g. omission), the neuronal activation of D2-MSN in the NAc by
269 error choices plays an important role in avoiding a bad option, and the neuronal activity of D2-MSN had a

270 positive correlation between before and after learning, individual D2-MSNs in the NAc could encode the value
271 information associated with the specific cue, in particular negative value. Moreover, given that D1-MSN plays
272 an important role in a reward learning and a reward approaching behavior (Hikida et al., 2010; Kravitz et al.,
273 2012; Macpherson and Hikida, 2018; Shin et al., 2018; Tai et al., 2012), the neuronal activation of D1-MSN by
274 positive outcome (e.g. reward) can contribute to choose the cue associated with positive outcome. The reason
275 we could not find any correlation in D1-MSN was attributed to the property of the VD-Avoid in which mice
276 could not associate the specific cue with the positive outcome because the cue associated with reward changed
277 in every trial in the VD-Avoid task.

278 Previous studies indicate that D2-MSN plays an important role in aversive learning (Danjo et al.,
279 2014; Hikida et al., 2010). However, classical paradigms trained mice not to take any action against the cue or
280 the context associated with an aversive stimulus. Therefore, it was difficult to distinguish the motor component
281 from an aversive information processing because D2-MSN is also involved in action and locomotion
282 (Bakhurin et al., 2020; Kravitz et al., 2010; Tecuapetla et al., 2016; Tsutsui-Kimura et al., 2017a; Yao et al.,
283 2021). Furthermore, although an electric shock or an air puff was often used as an aversive stimulus, those
284 stimuli are innately aversive and could be different from a negative prediction error (e.g. omission). Because
285 our paradigm shares the other components from the strategy, we think that it is easier to extract the component
286 of the strategy. Our previous data showed that perturbation of the D2 receptor signalling by knockout of D2L
287 interfered with visual discrimination learning (Morita et al., 2016). By combining the previous study and the
288 current study, we suggest that mice simultaneously learn to choose a cue associated with a reward and to avoid
289 a cue not associated with a reward in classical discrimination learning.

290 It is known that dopamine neurons encode a reward prediction error and the firing rate of dopamine
291 neurons decreases (e.g. dopamine dip) when the subjects did not receive the expected reward (Schultz et al.,
292 1997). We expect that the dopamine dip by a reward omission contributes to the strategy to avoid a bad option.
293 Activation of D2-MSN by disinhibition of D2 receptor could contribute to suppressing the action associated
294 with a negative outcome under discrimination learning. Interestingly, dopamine neurons are heterogeneous,
295 and reward prediction error activity can be observed in the NAcCore but not NAcShell (Iino et al., 2020; Jong
296 et al., 2019). We mainly targeted the NAcCore and revealed that chemogenetic and optogenetic suppression of

297 D2-MSN in the NAcCore is sufficient to decrease the performance in the VD-Avoid task. Although we cannot
298 exclude the possibility that the NAcShell also contributes to the strategy to avoid a bad option, we revealed
299 that at least D2-MSN in the NAcCore plays an important role in the strategy to avoid a bad option. Moreover,
300 it is reported that the dorsomedial striatum (DMS) plays an important role in reinforcement learning (Kravitz et
301 al., 2012; Tai et al., 2012) and electrophysiological recording showed that D1-MSN in the DMS preferentially
302 encodes positive prediction error and D2-MSN encodes negative prediction error (Shin et al., 2018). Therefore,
303 the DMS could also contribute to the strategy. In the future, we need to investigate how differently the
304 NAcCore, NAcShell, and DMS contribute to the execution of the strategy under decision.

305 Disability to execute the strategy to avoid a bad option could be associated with risk-taking behaviors.
306 A previous study showed D2-MSN activity in the NAc is important for avoiding risk-taking behavior
307 (Zalocusky et al., 2016). They showed D2-MSN activation in the NAc contributes to avoiding a risky choice.
308 Our results basically support their data and extend it to that D2-MSN activity in the NAc is involved in
309 suppressing the action associated with the negative outcome as well as a risky choice. However, they showed
310 D2-MSN activation occurs before the decision in the trial after the failure of risky-choice and optogenetic
311 activation of D2-MSN during the decision period is sufficient to avoid a risky choice. Our data showed that
312 optogenetic inhibition of D2-MSN in the NAc during ITI or Cue period did not affect the decision. Although
313 the exact mechanism is unknown because they did not test the effect of optogenetic activation of D2-MSN in
314 the NAc during outcome period on a risky choice, the timing of activation of D2-MSN in the NAc may differ
315 depending on the task. Moreover, their target is a medial region of the NAc, while our target is the lateral
316 region. Because dopamine dynamics are different depending on the region of the NAc (Iino et al., 2020; Jong
317 et al., 2019), the difference in activation timing may depend on the region. However, further investigations are
318 needed in the future.

319 Interestingly, our results showed that optogenetic inhibition of D2-MSN in the NAc during ITI or Cue
320 period did not affect the performance, suggesting that D2-MSN activity does not contribute to discrimination
321 itself. Those results also suggest that the action value information is transmitted from the NAc via the
322 thalamocortical loop and encoded in the neurons in different areas such as the parietal cortex (Akrami et al.,
323 2018) and the DMS (Tai et al., 2012) during the decision period. Moreover, because it is known that D2-MSN

324 in the NAc sends inhibitory input onto the VP (Gallo et al., 2018) and inhibition of GABAergic neurons in the
325 VP reduces the risky choice (Farrell et al., 2021), inhibition of GABAergic neurons in the VP via D2-MSN in
326 the NAc could be important for avoiding a bad option.

327 A previous study has shown that inhibition of D2-MSN in the NAc could lead to rewarding effects
328 (Gallo et al., 2018). However, our results cannot be explained by simple rewarding effects because optogenetic
329 inhibition of D2-MSN in the NAc selectively sped up the response latency in the trial after error. If the
330 rewarding effect by optogenetic inhibition affects the performance in the VD-Avoid task, the response latency
331 in the trial after correct should be similarly decreased. Given that activation of D2-MSN in the NAc increases
332 the motivation, D2-MSN in the NAc could be involved in decision making via complex mechanisms. This
333 history-dependent effect on response latencies is interesting from a psychological point of view. A number of
334 psychological studies has reported that subjects tend to slow down the response on the current trial after having
335 committed an error on the previous trial (Houtman et al., 2012; Notebaert et al., 2009; Rabbitt and Rodgers,
336 1977; Ruitenberg et al., 2014). Although the neural mechanisms underlying the post-error slowing have been
337 not fully understood, in this study we have shown that inhibition of D2-MSN in the NAc sped up the response
338 latency in the trials after error choice. The result suggests that activation of D2-MSN in the NAc after error
339 choice contributes to the post-error slowing.

340 Contradictory to a classical model, our results showed a subset of D1-MSNs in the NAc encodes the
341 negative information. However, given that it is reported that a subset of D1-MSN in the dorsal striatum
342 selectively encodes the aversive information (Kim et al., 2020; Xiao et al., 2020), the NAcCore could also
343 include the similar subset of D1-MSN. Given that majority of D1-MSN in the NAc were activated by reward
344 and bulk inhibition of D1-MSN decreased the motivation in our experiments, we conclude that D1-MSN in the
345 NAc mainly plays an important role in motivation. However, interestingly, chemogenetic suppression of D1-
346 MSN in the NAc affected the performance in a recent history-dependent manner. This result cannot be
347 explained by the simple effect of decreased motivation. Although further experiments are required, we think
348 that there are at least three subpopulations in D1-MSN of the NAc. The first population is involved in a general
349 motivation control. The second modulates the behavior based on reward history. The third encodes a negative
350 prediction error. Although we could not analyze the contribution of D1-MSN to the strategy because bulk

351 suppression of D1-MSN in the NAc drastically decreased the motivation, given that D1-MSN in the NAc plays
352 an important role in approaching behaviors in the Pavlovian conditioning (Macpherson and Hikida, 2018), we
353 expect that a subset of D1-MSN in the NAc contribute to the strategy to choose a good option. Considering
354 that D2-MSN in the NAc were activated by error and inhibition of D2-MSN in the NAc decreased the
355 performance in the trials after error response, D1-MSN and D2-MSN in the NAc have opposite functions in
356 the visual discrimination learning and activation of D1-MSN by reward and activation of D2-MSN by error
357 could lead to making the same choice (e.g. via high confidence) and avoid the previous choice, respectively.
358 Moreover, D2-MSN is involved in reversal learning and switching behaviors (Macpherson et al., 2016;
359 Matamales et al., 2020; Tsutsui-Kimura et al., 2017b; Yawata et al., 2012). Because these behaviors are
360 modulated by negative prediction error (e.g. reward omission), the strategy to avoid a bad option is required.
361 Although further experiments are needed to test if the same population regulate the strategy to avoid a bad
362 option and reversal learning, we conclude that D2-MSN in the NAc is important for detecting a negative
363 outcome and avoiding the action associated with the negative outcome.

364 It is known that decision strategy could be influenced by a brain state and maladaptive decision
365 making can be observed in various neurological and psychiatric disorders such as Parkinson's disease, drug
366 addiction, and mood disorders (Frank et al., 2004; Hikida et al., 2010; Kunisato et al., 2012; Strickland et al.,
367 2016). In Parkinson's disease, it is known that cell death of dopamine neurons occurs and acute dopamine
368 depletion leads to overactivation of D2-MSN (Parker et al., 2018). Moreover, a study shows that repeated
369 cocaine treatment reduced the frequency of miniature excitatory postsynaptic currents in D2-MSN (Kim et al.,
370 2011). These studies suggest that increase in excitability of D2-MSN leads to the strategy to avoid a bad
371 option, while decrease in excitability of D2-MSN causes to disability to avoid a bad option. These bidirectional
372 effects on the strategy support our hypothesis that activation of D2-MSN plays an important role in avoiding a
373 bad option. Although it is difficult to explain with one factor because the plasticity change occurs at neural
374 circuit level in various neurological and psychiatric disorders (Macpherson and Hikida, 2019), it is likely that
375 excitability of D2-MSN contributes to the strategy to avoid a bad option. In contrast to L-dopa treatment,
376 depression selectively disrupts the strategy to choose a good option (Kunisato et al., 2012). Because depression
377 is strongly related to the concentration of serotonin in the brain (Warden et al., 2012), the abnormal neural

activity of a subset of the neuron in the NAc via serotonin receptor could be associated with the disability to execute the strategy to choose a good option.

In conclusion, we provide the evidence in which D2-MSN in the NAc selectively contributes to the strategy to avoid a bad option. Moreover, activation of D2-MSN in the NAc by error choices plays an important role in acquisition of the strategy to avoid a bad option. Understanding the neural mechanisms underlying the strategy will help to treat with dysfunctions to execute the appropriate strategy in various psychiatric disorders such as Parkinson's disease, drug addiction, and mood disorders. In addition, the development of VD-Attend and VD-Avoid task could lead to a useful novel model to understand the neural mechanisms by which different people execute different strategies depending on the genes, situations, developments, and ages.

## ACKNOWLEDGEMENTS

This work was supported by MEXT/JSPS KAKENHI grants numbers, JP21K15209 (to T.N.), JP21K15210 (to T.M.), JP19H04983 and JP21H02804 (to K.H.), JP16H06568 and JP18H02542 (to T.H.), by AMED under grant number JP21wm0425010 (to T.H.), by Takeda Life Science Research Foundation (to T.H.), by the Salt Science Research Foundation (No. 2137 to T.H.), by Taiju Life Social Welfare Foundation (to T.H.), and by Smoking Research Foundation (to T.H.). We thank members of Hikida laboratory for maintenance of animals and helpful discussion.

## AUTHOR CONTRIBUTIONS

T.N. conceived the project, conducted all the experiments, performed the analysis, and wrote the manuscript. T.M. supported calcium imaging experiments. K.H. and T.H. contributed to the analysis. T.H. supervised the project. T.M., K.H., and T.H. reviewed and edited the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing financial interests.

405

406 **REFERENCES**

407


408 Aberman, J.E., Ward, S.J., and Salamone, J.D. (1998). Effects of Dopamine Antagonists and Accumbens
409 Dopamine Depletions on Time- Constrained Progressive-Ratio Performance. Pharmacology Biochemistry and
410 Behavior *61*, 341–348.

411 Akrami, A., Kopec, C.D., Diamond, M.E., and Brody, C.D. (2018). Posterior parietal cortex represents sensory
412 history and mediates its effects on behaviour. Nature *554*, 368–372.

413 Aragona, B.J., Liu, Y., Yu, Y.J., Curtis, J.T., Detwiler, J.M., Insel, T.R., and Wang, Z. (2006). Nucleus
414 accumbens dopamine differentially mediates the formation and maintenance of monogamous pair bonds. Nat
415 Neurosci *9*, 133–139.

416 Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., and Roth, B.L. (2007). Evolving the lock to fit the key to
417 create a family of G protein-coupled receptors potently activated by an inert ligand. Proc National Acad Sci
418 *104*, 5163–5168.

419 Bakhurin, K.I., Li, X., Friedman, A.D., Lusk, N.A., Watson, G.D., Kim, N., and Yin, H.H. (2020). Opponent
420 regulation of action performance and timing by striatonigral and striatopallidal pathways. Elife *9*, e54831.

421 Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience.
422 Psychopharmacology *191*, 391–431.

423 Berridge, K.C., and Kringelbach, M.L. (2008). Affective neuroscience of pleasure: reward in humans and
424 animals. Psychopharmacology *199*, 457–480.

425 Carlezon, W.A., and Thomas, M.J. (2009). Biological substrates of reward and aversion: A nucleus accumbens
426 activity hypothesis. Neuropharmacology *56*, 122–132.

427 Cepeda, C., Buchwald, N.A., and Levine, M.S. (1993). Neuromodulatory actions of dopamine in the
428 neostriatum are dependent upon the excitatory amino acid receptor subtypes activated. Proc National Acad Sci
429 *90*, 9576–9580.

430 Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for
431 reward and punishment in the ventral tegmental area. Nature *482*, 85–88.

432 Dana, H., Sun, Y., Mohar, B., Hulse, B.K., Kerlin, A.M., Hasseman, J.P., Tsegaye, G., Tsang, A., Wong, A.,
433 Patel, R., et al. (2019). High-performance calcium sensors for imaging activity in neuronal populations and
434 microcompartments. Nat Methods *16*, 649–657.

435 Danjo, T., Yoshimi, K., Funabiki, K., Yawata, S., and Nakanishi, S. (2014). Aversive behavior induced by
436 optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in
437 the nucleus accumbens. Proc National Acad Sci *111*, 6455–6460.

438 Delgado, M.R. (2007). Reward-Related Responses in the Human Striatum. Ann Ny Acad Sci *1104*, 70–88.

439 Farrell, M.R., Esteban, J.S.D., Faget, L., Floresco, S.B., Hnasko, T.S., and Mahler, S.V. (2021). Ventral
440 Pallidum GABA Neurons Mediate Motivation Underlying Risky Choice. J Neurosci *41*, 4500–4513.

441 Ferenczi, E.A., Zalocusky, K.A., Liston, C., Grosenick, L., Warden, M.R., Amatya, D., Katovich, K., Mehta,
442 H., Patenaude, B., Ramakrishnan, C., et al. (2016). Prefrontal cortical regulation of brainwide circuit dynamics
443 and reward-related behavior. Science *351*, aac9698.

444 Frank, M.J., Seeberger, L.C., and O'Reilly, R.C. (2004). By carrot or by stick: cognitive reinforcement
445 learning in parkinsonism. Science *306*, 1940–1943.

446 Frank, M.J., Samanta, J., Moustafa, A.A., and Sherman, S.J. (2007). Hold your horses: impulsivity, deep brain
447 stimulation, and medication in parkinsonism. Science *318*, 1309–1312.

448 Gallo, E.F., Meszaros, J., Sherman, J.D., Chohan, M.O., Teboul, E., Choi, C.S., Moore, H., Javitch, J.A., and
449 Kellendonk, C. (2018). Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory
450 transmission to the ventral pallidum. Nat Commun *9*, 1086.

451 Gerfen, C.R., and Surmeier, D.J. (2011). Modulation of Striatal Projection Systems by Dopamine. Annu Rev
452 Neurosci *34*, 441–466.

453 Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Weele, C.M.V., Kennedy, R.T.,
454 Aragona, B.J., and Berke, J.D. (2015). Mesolimbic dopamine signals the value of work. Nat Neurosci *19*, 117–
455 126.

456 Han, X., Chow, B.Y., Zhou, H., Klapoetke, N.C., Chuong, A., Rajimehr, R., Yang, A., Baratta, M.V., Winkle,
457 J., Desimone, R., et al. (2011). A High-Light Sensitivity Optical Neural Silencer: Development and
458 Application to Optogenetic Control of Non-Human Primate Cortex. Frontiers Syst Neurosci *5*, 18.

459 Hikida, T., Kimura, K., Wada, N., Funabiki, K., and Nakanishi, S. (2010). Distinct roles of synaptic
460 transmission in direct and indirect striatal pathways to reward and aversive behavior. Neuron *66*, 896–907.

461 Houtman, F., Castellar, E.N., and Notebaert, W. (2012). Orienting to errors with and without immediate
462 feedback. J Cogn Psychol *24*, 278–285.

463 Iino, Y., Sawada, T., Yamaguchi, K., Tajiri, M., Ishii, S., Kasai, H., and Yagishita, S. (2020). Dopamine D2
464 receptors in discrimination learning and spine enlargement. Nature *579*, 555–560.

465 Jong, J.W. de, Afjei, S.A., Dorocic, I.P., Peck, J.R., Liu, C., Kim, C.K., Tian, L., Deisseroth, K., and Lammel,
466 S. (2019). A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System.
467 Neuron *101*, 133–151.e7.

468 Kang, B.J., Song, S.S., Wen, L., Hong, K., Augustine, G.J., and Baik, J. (2017). Effect of optogenetic
469 manipulation of accumbal medium spiny neurons expressing dopamine D2 receptors in cocaine-induced
470 behavioral sensitization. Eur J Neurosci *46*, 2056–2066.

471 Kim, C.K., Sanchez, M.I., Hoerbelt, P., Fenno, L.E., Malenka, R.C., Deisseroth, K., and Ting, A.Y. (2020). A
472 Molecular Calcium Integrator Reveals a Striatal Cell Type Driving Aversion. Cell *183*, 2003-2019.e16.

473 Kim, J., Park, B.-H., Lee, J.H., Park, S.K., and Kim, J.-H. (2011). Cell Type-Specific Alterations in the
474 Nucleus Accumbens by Repeated Exposures to Cocaine. Biol Psychiat *69*, 1026–1034.

475 Kravitz, A.V., Freeze, B.S., Parker, P.R.L., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010).
476 Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature *466*,
477 622–626.

478 Kravitz, A.V., Tye, L.D., and Kreitzer, A.C. (2012). Distinct roles for direct and indirect pathway striatal
479 neurons in reinforcement. Nat Neurosci *15*, 816–818.

480 Kunisato, Y., Okamoto, Y., Ueda, K., Onoda, K., Okada, G., Yoshimura, S., Suzuki, S., Samejima, K., and
481 Yamawaki, S. (2012). Effects of depression on reward-based decision making and variability of action in
482 probabilistic learning. J Behav Ther Exp Psy *43*, 1088–1094.

483 Lee, D. (2013). Decision Making: From Neuroscience to Psychiatry. Neuron *78*, 233–248.

484 Lobo, M.K., Covington, H.E., Chaudhury, D., Friedman, A.K., Sun, H., Damez-Werno, D., Dietz, D.M.,
485 Zaman, S., Koo, J.W., Kennedy, P.J., et al. (2010). Cell Type–Specific Loss of BDNF Signaling Mimics
486 Optogenetic Control of Cocaine Reward. Science *330*, 385–390.

487 Lu, X.-Y., Ghasemzadeh, M.B., and Kalivas, P.W. (1997). Expression of D1 receptor, D2 receptor, substance
488 P and enkephalin messenger RNAs in the neurons projecting from the nucleus accumbens. Neuroscience *82*,
489 767–780.

490 Luo, Y.-J., Li, Y.-D., Wang, L., Yang, S.-R., Yuan, X.-S., Wang, J., Cherasse, Y., Lazarus, M., Chen, J.-F.,
491 Qu, W.-M., et al. (2018). Nucleus accumbens controls wakefulness by a subpopulation of neurons expressing
492 dopamine D1 receptors. Nat Commun *9*, 1576.

493 Macpherson, T., and Hikida, T. (2018). Nucleus Accumbens Dopamine D1-Receptor-Expressing Neurons
494 Control the Acquisition of Sign-Tracking to Conditioned Cues in Mice. Front Neurosci-Switz *12*, 418.

495 Macpherson, T., and Hikida, T. (2019). Role of basal ganglia neurocircuitry in the pathology of psychiatric
496 disorders. Psychiat Clin Neuros *73*, 289–301.

497 Macpherson, T., Morita, M., Wang, Y., Sasaoka, T., Sawa, A., and Hikida, T. (2016). Nucleus accumbens
498 dopamine D2-receptor expressing neurons control behavioral flexibility in a place discrimination task in the
499 IntelliCage. Learn Memory *23*, 359–364.

500 Maia, T.V., and Frank, M.J. (2011). From reinforcement learning models to psychiatric and neurological
501 disorders. Nat Neurosci *14*, 154–162.

502 Matamales, M., McGovern, A.E., Mi, J.D., Mazzone, S.B., Balleine, B.W., and Bertran-Gonzalez, J. (2020).
503 Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. Science *367*, 549–555.

504 Morita, M., Wang, Y., Sasaoka, T., Okada, K., Niwa, M., Sawa, A., and Hikida, T. (2016). Dopamine D2L
505 Receptor Is Required for Visual Discrimination and Reversal Learning. Mol Neuropsychiatry *2*, 124–132.

506 Notebaert, W., Houtman, F., Opstal, F.V., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing:
507 An orienting account. Cognition *111*, 275–279.

508 Parker, J.G., Marshall, J.D., Ahanonu, B., Wu, Y.-W., Kim, T.H., Grewe, B.F., Zhang, Y., Li, J.Z., Ding, J.B.,
509 Ehlers, M.D., et al. (2018). Diametric neural ensemble dynamics in parkinsonian and dyskinetic states. Nature
510 *557*, 177–182.

Parker, N.F., Cameron, C.M., Taliaferro, J.P., Lee, J., Choi, J.Y., Davidson, T.J., Daw, N.D., and Witten, I.B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat Neurosci *19*, 845–854.

Rabbitt, P., and Rodgers, B. (1977). What does a Man do after he Makes an Error? An Analysis of Response Programming. Q J Exp Psychol *29*, 727–743.

Roitman, M.F., Wheeler, R.A., and Carelli, R.M. (2005). Nucleus Accumbens Neurons Are Innately Tuned for Rewarding and Aversive Taste Stimuli, Encode Their Predictors, and Are Linked to Motor Output. Neuron *45*, 587–597.

Roitman, M.F., Wheeler, R.A., Wightman, R.M., and Carelli, R.M. (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. Nat Neurosci *11*, 1376–1377.

Ruitenberg, M.F.L., Abrahamse, E.L., Kleine, E.D., and Verwey, W.B. (2014). Post-error slowing in sequential action: an aging study. Front Psychol *5*, 119.

Salamone, J.D., Correa, M., Farrar, A., and Mingote, S.M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. Psychopharmacology *191*, 461–482.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. Science *310*, 1337–1340.

Schultz, W., Apicella, P., Scarnati, ”  Eugenio , and Ljungbergc, T. (1992). Neuronal Activity in Monkey Ventral Striatum Related to the Expectation of Reward. The Journal of Neuroscience *12*, 4595–4610.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

Sheintuch, L., Rubin, A., Brande-Eilat, N., Geva, N., Sadeh, N., Pinchasof, O., and Ziv, Y. (2017). Tracking the Same Neurons across Multiple Days in Ca2+ Imaging Data. Cell Reports *21*, 1102–1115.

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D.J. (2008). Dichotomous Dopaminergic Control of Striatal Synaptic Plasticity. Science *321*, 848–851.

Shin, J.H., Kim, D., and Jung, M.W. (2018). Differential coding of reward and movement information in the dorsomedial striatal direct and indirect pathways. Nat Commun *9*, 404.

Stephenson-Jones, M., Yu, K., Ahrens, S., Tucciarone, J.M., Huijstee, A.N. van, Mejia, L.A., Penzo, M.A., Tai, L.-H., Wilbrecht, L., and Li, B. (2016). A basal ganglia circuit for evaluating action outcomes. Nature *539*, 289–293.

Stoof, J.C., and Kebabian, J.W. (1984). Two dopamine receptors: Biochemistry, physiology and pharmacology. Life Sci *35*, 2281–2296.

Strickland, J.C., Bolin, B.L., Lile, J.A., Rush, C.R., and Stoops, W.W. (2016). Differential sensitivity to learning from positive and negative outcomes in cocaine users. Drug Alcohol Depen *166*, 61–68.

Surmeier, D.J., Ding, J., Day, M., Wang, Z., and Shen, W. (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. Trends Neurosci *30*, 228–235.

546 Surmeier, D.J., Plotkin, J., and Shen, W. (2009). Dopamine and synaptic plasticity in dorsal striatal circuits
547 controlling action selection. Curr Opin Neurobiol *19*, 621–628.

548 Tai, L.-H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct
549 subpopulations of striatal neurons mimics changes in action value. Nat Neurosci *15*, 1281–1289.

550 Tecuapetla, F., Jin, X., Lima, S.Q., and Costa, R.M. (2016). Complementary Contributions of Striatal
551 Projection Pathways to Action Initiation and Execution. Cell *166*, 703–715.

552 Tsutsui-Kimura, I., Takiue, H., Yoshida, K., Xu, M., Yano, R., Ohta, H., Nishida, H., Bouchekioua, Y.,
553 Okano, H., Uchigashima, M., et al. (2017a). Dysfunction of ventrolateral striatal dopamine receptor type 2-
554 expressing medium spiny neurons impairs instrumental motivation. Nat Commun *8*, 14304.

555 Tsutsui-Kimura, I., Natsubori, A., Mori, M., Kobayashi, K., Drew, M.R., d'Exaerde, A. de K., Mimura, M.,
556 and Tanaka, K.F. (2017b). Distinct Roles of Ventromedial versus Ventrolateral Striatal Medium Spiny
557 Neurons in Reward-Oriented Behavior. Curr Biol *27*, 3042-3048.e4.

558 Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science *211*,
559 453–458.

560 Warden, M.R., Selimbeyoglu, A., Mirzabekov, J.J., Lo, M., Thompson, K.R., Kim, S.-Y., Adhikari, A., Tye,
561 K.M., Frank, L.M., and Deisseroth, K. (2012). A prefrontal cortex-brainstem neuronal projection that controls
562 response to behavioural challenge. Nature *492*, 428–432.

563 Wulff, P., and Arenkiel, B.R. (2012). Chemical genetics: receptor–ligand pairs for rapid manipulation of
564 neuronal activity. Curr Opin Neurobiol *22*, 54–60.

565 Xiao, X., Deng, H., Furlan, A., Yang, T., Zhang, X., Hwang, G.-R., Tucciarone, J., Wu, P., He, M.,
566 Palaniswamy, R., et al. (2020). A Genetically Defined Compartmentalized Striatal Direct Pathway for
567 Negative Reinforcement. Cell *183*, 211-227.e20.

568 Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G.C.R., Urakubo, H., Ishii, S., and Kasai, H. (2014). A critical
569 time window for dopamine actions on the structural plasticity of dendritic spines. Science *345*, 1616–1620.

570 Yao, Y., Gao, G., Liu, K., Shi, X., Cheng, M., Xiong, Y., and Song, S. (2021). Projections from D2 Neurons in
571 Different Subregions of Nucleus Accumbens Shell to Ventral Pallidum Play Distinct Roles in Reward and
572 Aversion. Neurosci Bull *37*, 623–640.

573 Yawata, S., Yamaguchi, T., Danjo, T., Hikida, T., and Nakanishi, S. (2012). Pathway-specific control of
574 reward learning and its flexibility via selective dopamine receptors in the nucleus accumbens. Proc National
575 Acad Sci *109*, 12764–12769.

576 Zalocusky, K.A., Ramakrishnan, C., Lerner, T.N., Davidson, T.J., Knutson, B., and Deisseroth, K. (2016).
577 Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. Nature *531*, 642–646.

578 Zhou, L., Furuta, T., and Kaneko, T. (2003). Chemical organization of projection neurons in the rat accumbens
579 nucleus and olfactory tubercle. Neuroscience *120*, 783–798.

580 Zhou, P., Resendez, S.L., Rodriguez-Romaguera, J., Jimenez, J.C., Neufeld, S.Q., Giovannucci, A., Friedrich,
581 J., Pnevmatikakis, E.A., Stuber, G.D., Hen, R., et al. (2018). Efficient and accurate extraction of in vivo
582 calcium signals from microendoscopic video data. Elife *7*, e28728.

583

584

585 **METHOD DETAILS**

586 **Animals**

587 Wild-type C57BL/6J mice (male, 8-10 weeks old) were used for validation of behavioral experiments. For
588 DREADD, optogenetic, and calcium imaging experiments, male heterozygous D1-Cre (FK150Gsat) and D2-
589 Cre (ER44Gsat) mice were used (DREADD, D1-Cre, n = 12, one mouse was excluded due to no viral
590 expression; D2-Cre, n = 12, one mouse was excluded due to unstable behavior; optogenetics, D2-Cre, n = 19, 2
591 mice were excluded due to insufficient conditioning; calcium imaging, D1-Cre, n = 3; D2-Cre, n = 4, one
592 mouse was excluded because GRIN lens was located anterior to the NAc). D1-Cre and D2-Cre were
593 maintained in a C57BL/6J background. Animals were housed on a 12-hour light/dark cycle. Behavioral studies
594 were conducted during the light cycle. Mice were kept on water restriction during behavioral testing. For all
595 behavioral experiments except for calcium imaging experiment, mice were grouped housed throughout the
596 experiments. For calcium imaging experiment, mice were singly housed after GRIN lens implantation. All
597 experiments conformed to the guidelines of the National Institutes of Health experimental procedures, and
598 were approved by the Animal Experimental Committee of Institute for Protein Research at Osaka University
599 (approval ID 29-02-1).

600

601 **Stereotaxic Surgery**

602 All mice used in this study were anesthetized with ketamine (100 mg/kg) and xylazine (20 mg/kg) for surgical
603 procedures and placed in a stereotaxic frame (Kopf Instruments).

604 For DREADD experiments, heterozygous D1-Cre or D2-Cre mice were bilaterally injected with 400
605 nl of AAV5-hSyn-DIO-hM4Di(Gi)-mCherry ($5.1\times10^{12}$ GC/ml, UNC) or AAV5-hSyn-DIO-mCherry ($5.2\times10^{12}$
606 GC/ml, UNC) using a Nanoject II instrument (Drummond) at a rate of 100 nl/min (coordinates in mm: AP
607 +1.20, ML ±1.25 from bregma, and DV −3.50 from brain surface. The injection pipette remained in place for
608 5–10 min to reduce backflow.

609 For optogenetics experiments, heterozygous D2-Cre mice were bilaterally injected with 400 nl of
610 AAV5-CAG-FLEX-ArchT3.0-tdTomato ($1.3\times10^{13}$ GC/ml, Addgene) or AAV5-EF1a-DIO-eYFP ($1.3\times10^{13}$
611 GC/ml, Addgene) were using a Nanoject III instrument (Drummond) at a rate of 100 nl/min (coordinates in
612 mm: AP +1.20, ML ±1.25 from bregma, and DV −3.50 from brain surface. The injection pipette remained in
613 place for 5–10 min to reduce backflow. After retraction, 200 µm diameter (NA 0.37) optic fibers (Thorlabs)
614 were bilaterally implanted with the dental cement (Superbond) at AP +1.20, ML ±1.30 from bregma, and DV
615 −3.20 from brain surface.
616 For calcium imaging experiments, heterozygous D1-Cre or D2-Cre mice were unilaterally injected
617 with 1200 nl of AAV9-FLEX-jGCaMP7f ($9.6\times10^{12}$ GC/ml, Addgene) were stereotaxically injected using a
618 Nanoject III instrument (Drummond) at a rate of 100 nl/min (coordinates in mm: AP +1.20, ML ±1.25 from
619 bregma, and DV −3.60 and −3.10 from brain surface. The injection pipette remained in place for 5–10 min to
620 reduce backflow. After virus injection, a sterile 21-gauge needle was slowly lowered into the brain to a depth
621 of –2.0 mm from the brain surface to aspirate brain tissue above the NAc. The GRIN lens (600 µm diameter,
622 Inscopix) was slowly lowered into the brain to a depth of –3.20 mm from the brain surface by using a GRIN
623 lens holder (Inscopix). We secured the GRIN lens to the skull with the dental cement (Superbond). A silicone
624 elastomer (Kwik-Cast; World Precision Instruments) was applied to the top of the lens to prevent an external
625 damage. Four to six weeks after lens implantation, a baseplate (Inscopix) attached to the miniature microscope
626 (nVista; Inscopix) was positioned above the GRIN lens. The focal plane was adjusted on the basis of the
627 position that blood vessels were clearly observed. After positioning adjustment, the baseplate was secured with
628 the dental cement.
629
630 **Behavioral Experiments**
631 *Apparatus.* Training and testing were conducted in a Bussey-Saksida touchscreen chamber (Lafayette
632 Instrument). A black plastic mask with 2 windows (70×75 $mm^2$ spaced, 5 mm apart, 16 mm above the floor)
633 was placed in front of the touchscreen. ABET II and WhiskerServer software (Lafayette) were used to control
634 operant system and data collection.
635

636 *Pretraining.* As the first phase (3 days), mice were habituated to the chamber in 40-min sessions. Diluted
637 condensed milk (7 µl, Morinaga Milk) was dispensed in the food magazine every 10 sec. In the following
638 phase (1 day), a stimulus was randomly displayed in one of the 2 windows. After a 30-sec stimulus
639 presentation, the milk reward (20 µl) was delivered with a tone (3 kHz) and magazine light. When mice
640 collected the reward, the magazine light went out, and the next trial commenced (60 trials, or up to 60 min)
641 with a new stimulus after a 20-sec intertrial interval (ITI). In the next phase, stimuli were randomly displayed
642 in one of 2 windows, and mice were obligated to touch the stimulus to receive a reward. In the final phase of
643 the pretraining, when a blank window was touched, mice were punished with a 5-sec time-out. After reaching
644 criterion (77% correct for 2 consecutive days), mice moved on to basic training.
645
646 *Basic training.* Mice were tested 5–6 days per week (60 trials per day, or up to 60 min). Each trial
647 was initiated after mice nose-poked in the magazine. The visual cues were presented until mice respond to any
648 windows.
649 For VD-Attend task, two visual cues (Marble and a random image) were presented in the touchscreen.
650 A random image was pseudorandomly chosen from 51 images. If a mouse responded to visual cue "Marble",
651 milk reward (7 µl) was delivered with a tone (3 kHz) and magazine light. When mice collected the reward, the
652 magazine light went out, and the next trial commenced (60 trials, or up to 60 min) with a new stimulus after a
653 20-sec intertrial interval (ITI). If a mouse responded to visual cue "a random image", the mouse was punished
654 with a 5-sec time-out (house light on).
655 For VD-avoid task, two visual cues (Flag and a random image) were presented in the touchscreen. If a
656 mouse responded to visual cue "a random image", milk reward (7 µl) was delivered with a tone (3 kHz) and
657 magazine light. If a mouse responded to visual cue "Flag", the mouse was punished with a 5-sec time-out
658 (house light on).
659 A response to a random image was recorded as a correct response, while a response to visual cue "Flag" was
660 recorded as an incorrect response.
661 After reaching criterion (>80% correct for 2 consecutive days), mice moved on to probe test for
662 DREADD experiments, cable habituation for optogenetic experiments, respectively.

663 For DREADD experiments, vehicle at day1, 2, 3, 5, or CNO (1.0 mg/kg diluted with vehicle, Sigma
664 Aldrich) at day4, 6 was intraperitoneally administered 30 min before the session.
665 For optogenetic inhibition experiments, once the performance stabilized (>80% correct for 2
666 consecutive days) with cable, optogenetic stimulation starts. LED power was set to 1-3 mW. Stimulation
667 schedule was counterbalanced.
668 For calcium imaging experiments, data was acquired at 20 Hz with 0.6 mW LED at the first session of
669 the basic training (Novice) and the session after reaching criterion (the criterion session; Expert). After
670 acquisition, calcium recording files were temporally (factor of 2) and spatially (factor of 4) downsampled and
671 motion-corrected using Inscopix Data Processing software ver 1.3.0. The fluorescent traces of individual
672 neurons were extracted from these images by CNMFe (Zhou et al., 2018). Z-scores were calculated from all
673 recording data.
674 Percentage correct (correct trials divided by correct plus incorrect trials, recorded as percent), and
675 latencies to correct response, incorrect response, and reward collection were monitored in all behavioral
676 experiments.
677
678 **Histology**
679 Animals were deeply anesthetized and transcardially perfused with 0.01 M PBS followed by 4%
680 paraformaldehyde (PFA) in 0.1 M PB (pH 7.4). Brains were removed and post-fixed with 4% PFA at 4 ºC for
681 2 days. After cryoprotection, brains were embedded in OCT compound and cryosectioned (40 µm). Sections
682 were mounted with antifade mouting medium with DAPI (Vectashield). Stitched images were acquired using a
683 Keyence BZ-X800 microscope.
684
685 **Statistical analyses**
686 Prism (Graphpad) software was used for statistical analyses. The behavioral performances in wild-type were
687 analyzed using unpaired t-test. DREADD data were analyzed using two-way RM ANOVA with Group
688 (hM4Di, mCherry) and Drug Treatment (vehicle, CNO) or one-way ANOVA (D1-hM4Di, D2-hM4Di, D1/D2-
689 mCherry). Optogenetic data were analyzed using two-way RM ANOVA with Group (ArchT, eYFP) and Light

690 stimulation (OFF, ON) or History (After Correct, After Error) and Light stimulation (OFF, ON). Post hoc
691 Sidak's multiple comparisons test was performed when F-ratios were significant ($p < 0.05$). Comparisons of
692 proportion of cells were made using chi-squared test. Simple linear regression was made to calculate the
693 correlation of auROC between novice and expert. Frequency distributions were compared using the
694 Kolmogorov-Smirnov (KS) test. All data are expressed as means ± SEM.
695
696
697 **Figure Legends**
698 **Figure 1. Experimental Design and Behavioral Performances**
699 (A) Experimental design.
700 (B) Timeline of the task events and the definition of the behavioral parameters.
701 (C) The percentage of correct in each session (VD-Attend, n = 11; VD-Avoid, n = 12).
702 (D) The number of sessions to criterion of the VD-Avoid was higher than that of the VD-Attend task (unpaired
703 t-test, $t_{21} = 2.413$, *p = 0.0251).
704 (E) Total errors to criterion of the VD-Avoid were higher than that of the VD-Attend task (unpaired t-test, $t_{21}$ =
705 2.288, *p = 0.0326).
706 (F and G) The behavioral performances were similar between the VD-Attend and VD-Avoid task in the 1$^{st}$
707 session (F, unpaired t-test, $t_{21} = 0.7053$, $p = 0.4884$) and the last session (G, unpaired t-test, $t_{21} = 1.613$, p =
708 0.1217).
709 (H and I) Correct latencies were similar between the VD-Attend and VD-Avoid task in the last session (H,
710 unpaired t-test, $t_{21} = 1.420$, $p = 0.1704$). Reward latencies were similar between the VD-Attend and VD-Avoid
711 task in the last session (I, unpaired t-test, $t_{21} = 0.007948$, $p = 0.9937$).
712
713 **Figure 2. DREADD Suppression of D2-MSN in the NAc Selectively Decreases the Performance of the**
714 **VD-Avoid Task**
715 (A) Experimental timeline.

716 (B) Representative injection site of AAV5-hSyn-DIO-hM4Di-mCherry (Left). Scale bar, 500 μm. NAcCore,
717 nucleus accumbens core; NAcSh, nucleus accumbens shell. Drawings of superimposed AAV injection sites in
718 the NAc (Right).
719 (C) The effects of DREADD suppression on the behavioral performance of each task in D1-Cre mice (Left, n
720 = 7), D2-Cre mice (Middle, n = 8). The effects of CNO treatment on the behavioral performance of each task
721 in D1/D2-mCherry mice (Right, n = 7).
722 (D) DREADD suppression of D1-MSN in the NAc decreased the behavioral performance in both the VD-
723 Attend and VD-Avoid task (Left, Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,12}$ =
724 19.83, ***p = 0.0008; VD-Attend, **p = 0.0071; VD-Avoid, *p = 0.0394). DREADD suppression of D2-MSN
725 in the NAc selectively decreased the behavioral performance in the VD-Avoid but not VD-Attend task
726 (Middle, Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,14}$ = 40.88, ***p < 0.0001; VD-
727 Attend, p = 0.2334; VD-Avoid, ***p < 0.0001). CNO treatment in D1/D2-mCherry did not affect the
728 behavioral performance in both the VD-Attend and VD-Avoid task (Right, Two-way RM-ANOVA with Sidak
729 correction, Treatment effects, $F_{1,12}$ = 0.8707, p = 0.3692; VD-Attend, p = 0.7590; VD-Avoid, p = 0.1326).
730 (E) The effects of DREADD suppression on the behavioral performance of the VD-Attend task across all
731 groups (Top, One-way ANOVA with Sidak correction, Group effects, $F_{2,19}$ = 8.688, **p = 0.0021; D1-hM4Di
732 vs D2-hM4Di, **p = 0.0078; D1-hM4Di vs D1/D2-mCherry, **p = 0.0039; D2-hM4Di vs D1/D2-mCherry, p
733 = 0.9643). The effects of DREADD suppression on the behavioral performance of the VD-Avoid task across
734 all groups (Bottom, One-way ANOVA with Sidak correction, Group effects, $F_{2,19}$ = 9.396, **p = 0.0015; D1-
735 hM4Di vs D2-hM4Di, p = 0.9999; D1-hM4Di vs D1/D2-mCherry, **p = 0.0041; D2-hM4Di vs D1/D2-
736 mCherry, **p = 0.0036).
737 (F and G) The effects of DREADD suppression on the motivation indices. DREADD suppression of D1-MSN
738 in the NAc decreased the number of earned reward (F) in both the VD-Attend and VD-Avoid task (Left, Two-
739 way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,12}$ = 70.32, ***p < 0.0001; VD-Attend, ***p =
740 0.0001; VD-Avoid, ***p = 0.0002). DREADD suppression of D2-MSN in the Nac did not affect the number
741 of earned reward in both the VD-Attend and VD-Avoid task (Middle, Two-way RM-ANOVA with Sidak
742 correction, Treatment effects, $F_{1,14}$ = 0.2215, p = 0.6451; VD-Attend, p = 0.7989; VD-Avoid, p = 0.3959).

743 CNO treatment in D1/D2-mCherry did not affect the number of earned reward in both the VD-Attend and VD-
744 Avoid task (Right, Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,12} = 0.6391$, p =
745 0.4396; VD-Attend, p = 0.9626; VD-Avoid, p = 0.1759). DREADD suppression of D1-MSN in the Nac
746 increased reward latencies (G) in both the VD-Attend and VD-Avoid task (Left, Two-way RM-ANOVA,
747 Treatment effects, $F_{1,12} = 6.259$, *p = 0.0278). DREADD suppression of D2-MSN in the Nac did not affect
748 reward latencies in both the VD-Attend and VD-Avoid task (Middle, Two-way RM-ANOVA, Treatment
749 effects, $F_{1,14} = 0.8027$, p = 0.3854). CNO treatment in D1/D2-mCherry did not affect reward latencies in both
750 the VD-Attend and VD-Avoid task (Right, Two-way RM-ANOVA, Treatment effects, $F_{1,12} = 3.705$, p =
751 0.0783).
752
753 **Figure 3. Post-Error Activation of D2-MSN is Necessary for Avoiding a Bad Option**
754 (A) Schematic of viral injection and optic fiber implantation.
755 (B) Representative coronal section with optic fibers. Scale bar, 1 mm.
756 I Schematic of optical stimulation protocol.
757 (D) Optical suppression of D2-MSN in the NAc during ITI period did not affect the behavioral performance
758 (Two-way RM-ANOVA with Sidak correction, Group × Treatment interaction, $F_{1,15} = 0.05089$, p = 0.8246;
759 ArchT, n = 9, p = 0.9121; eYFP, n = 8, p = 0.9981).
760 (E) Optical suppression of D2-MSN in the NAc during Cue period did not affect the behavioral performance
761 (Two-way RM-ANOVA with Sidak correction, Group × Treatment interaction, $F_{1,15} = 0.8700$, p = 0.3657;
762 ArchT, p = 0.7815; eYFP, p = 0.7571).
763 (F) Optical suppression of D2-MSN in the NAc during Outcome period decreased the behavioral performance
764 in the next trial (Two-way RM-ANOVA with Sidak correction, Group × Treatment interaction, $F_{1,15} = 5.340$,
765 *p = 0.0355; ArchT, p = 0.0271; eYFP, p = 0.8368).
766 (G and H) Optical stimulation of D2-MSN in the NAc during Outcome period after error response decreased
767 the behavioral performance of ArchT mice (G) in the next trial (Two-way RM-ANOVA with Sidak correction,
768 Treatment effects, $F_{1,8} = 7.134$, *p = 0.0283; After Correct, p = 0.2634; After Error, **p = 0.0029) but not

769 eYFP mice (H, Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,7} = 0.2682$, $p = 0.6205$;

770 After Correct, $p = 0.9909$; After Error, $p = 0.8477$).

771

772 **Figure 4. Individual D2-MSNs Are Predominantly Activated by Error Choice**

773 (A) A schematic of viral injection and GRIN lens implantation.

774 (B) Representative coronal image of jGCaMP7f expression in D2-MSNs. Scale bar, 500 µm.

775 (C) Maximum projection image of a representative imaging plane. Scale bar, 50 µm.

776 (D) Example traces of individual neurons from a representative mouse performing the VD-Avoid task.

777 (E) Averaged traces of an example neuron showing error response. Trial-by-trial (Top) and averaged (Middle)

778 responses of an example neuron in correct and error trials. Area under the ROC curve (auROC) reflecting

779 statistical difference between calcium trace of error and correct trials (Bottom).

780 (F) Averaged traces of an example neuron showing correct response. Trial-by-trial (Top) and averaged

781 (Middle) responses of an example neuron in correct and error trials. Area under the ROC (auROC) curves

782 reflecting statistical difference between calcium trace of error and correct trials (Bottom).

783 (G) The first two components of the auROC curves of all the neurons of D1-Cre (n = 3 mice) and D2-Cre mice

784 (n = 3 mice).

785 (H) Heatmaps of the auROC curves of all the neurons of D1-Cre and D2-Cre mice. Each row represents one

786 neuron.

787 (I) Averaged traces of the auROC curves from the responsive Type I neurons (n = 78) or the responsive Type

788 II neurons (n = 53).

789 (J) The proportion of responsive cells across the session.

790 (K) Averaged proportion of responsive cells during ITI, Cue, or Outcome period. The proportion of responsive

791 cells during Outcome period was more than that during ITI or Cue period in both D1-Cre (Top, Chi-square test

792 with Bonferroni correction, $\chi^2 = 47.92$, ITI vs Cue, $p > 0.9999$; ITI vs Outcome, ***$p < 0.0001$; Cue vs

793 Outcome, ***$p < 0.0001$) and D2-Cre (Bottom, Chi-square test with Bonferroni correction, $\chi^2 = 40.37$, ITI vs

794 Cue, $p > 0.9999$; ITI vs Outcome, ***$p < 0.0001$; Cue vs Outcome, ***$p < 0.0001$) mice.

795 (L) The fraction of Type I and Type II neurons in D1-Cre or D2-Cre mice. The fraction of Type II neurons in
796 D2-Cre mice were significantly more than that in D1-Cre mice (Chi-square test, $\chi^2 = 10.38$, **p = 0.0013).
797
798 **Figure 5. More Non-Responsive D2-MSNs Acquire Error Responses than D1-MSNs through Learning**
799 (A) Representative image of contour map from novice (green) and expert (red).
800 (B) Averaged traces of an example neuron acquiring error response through learning. Averaged calcium traces
801 of novices and experts in correct and error trials (Top). Area under the ROC curve (auROC) reflecting
802 statistical difference between calcium trace of error and correct trials from novices and experts (Bottom).
803 (C) Averaged traces of an example neuron acquiring correct response through learning. Averaged calcium
804 traces of novices and experts in correct and error trials (Top). Area under the ROC curve (auROC) reflecting
805 statistical difference between calcium trace of error and correct trials from novices and experts (Bottom).
806 (D) Relationship between averaged auROC of novice and that of expert in D1-Cre (Left, n = 3 mice) or
807 D2-Cre mice (Middle, n = 3 mice). Schematic of classification of neuron type (Right). I, Type I; II, Type II;
808 N, Non-responsive.
809 (E) The fraction of each neuron type in D1-Cre (Left) and D2-Cre mice (Right).
810 (F) Non-responsive D2-MSN preferentially became Type I (Error Type) through learning (Left, Chi-square
811 test, $\chi^2 = 4.735$, *p = 0.0296), while Non-responsive D1-MSN preferentially became Type II (Correct Type)
812 through learning (Right, Chi-square test, $\chi^2 = 6.105$, *p = 0.0135).
813
814 **Supplementary Figure S1. Response Latencies Were Modulated by Recent History**
815 (A) The effects of recent history on response latencies in the VD-Attend task (unpaired t-test, $t_{10} = 0.8115$, p =
816 0.4359).
817 (B) The effects of recent history on response latencies in the VD-Avoid task (unpaired t-test, $t_{11} = 2.090$, p =
818 0.0607).
819
820 **Supplementary Figure S2. Cell Type Specific Gene Expression**

821 Representative coronal sections of DREADD expression in the D1-MSN or D2-MSN of the NAc (Left).
822 Representative coronal sections containing projection areas, ventral pallidum (VP, Middle) and substantia
823 nigra pars reticulata (SNr, Right).
824
825 **Supplementary Figure S3. DREADD Suppression of D1-MSN in the NAc Decreases the Behavioral**
826 **Performance in a Recent History Dependent Manner**
827 (A) DREADD suppression of D1-MSN in the NAc decreased the behavioral performance after correct trial in
828 the VD-Attend task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction, $F_{1,11}$ =
829 8.323, *p = 0.0148; After Correct, **p = 0.0015; After Error, p = 0.6357).
830 (B) DREADD suppression of D1-MSN in the NAc decreased the behavioral performance after correct trial in
831 the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction, $F_{1,12}$ =
832 5.603, *p = 0.0356; After Correct, *p = 0.0133; After Error, p = 0.9990).
833 (C) DREADD suppression of D2-MSN in the NAc did not affect the behavioral performance after correct or
834 error trial in the VD-Attend task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment
835 interaction, $F_{1,14}$ = 0.3984, p = 0.5381; After Correct, p = 0.7098; After Error, p = 0.3406).
836 (D) DREADD suppression of D2-MSN in the NAc decreased the behavioral performance after correct and
837 error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,14}$ =
838 29.79, ***p < 0.0001; After Correct, **p = 0.0012; After Error, **p = 0.0014).
839 (E) CNO treatment in D1/D2-mCherry did not affect the behavioral performance after correct or error trial in
840 the VD-Attend task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction, $F_{1,12}$ =
841 1.058, p = 0.3239; After Correct, p = 0.8609; After Error, p = 0.6787).
842 (F) CNO treatment in D1/D2-mCherry did not affect the behavioral performance after correct or error trial in
843 the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction, $F_{1,12}$ =
844 0.001573, p = 0.9690; After Correct, p = 0.7057; After Error, p = 0.7319).
845
846 **Supplementary Figure S4. Optic Fiber Placement**
847 Histology on optic fiber placement.

848

849 **Supplementary Figure S5. Optogenetic Suppression of D2-MSN in the NAc During ITI or Cue Period**

850 **Does Not affect the Behavioral Performance in a Recent History Dependent Manner**

851 (A) Optogenetic suppression of D2-MSN in the Nac during ITI period did not affect the behavioral

852 performance after correct or error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction,

853 Treatment effects, $F_{1,8} = 1.015$, $p = 0.3433$; After Correct, $p = 0.9990$; After Error, $p = 0.3294$).

854 (B) Light stimulation in D2-eYFP during ITI period did not affect the behavioral performance after correct or

855 error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment

856 interaction, $F_{1,7} = 2.535$, $p = 0.1554$; After Correct, $p = 0.9265$; After Error, $p = 0.1910$).

857 (C) Optogenetic suppression of D2-MSN in the NAc during Cue period did not affect the behavioral

858 performance after correct or error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction,

859 Treatment effects, $F_{1,8} = 0.7479$, $p = 0.4123$; After Correct, $p = 0.9760$; After Error, $p = 0.3471$).

860 (D) Light stimulation in D2-eYFP during Cue period did not affect the behavioral performance after correct or

861 error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment

862 interaction, $F_{1,7} = 1.224$, $p = 0.3051$; After Correct, $p = 0.9961$; After Error, $p = 0.3302$).

863

864 **Supplementary Figure S6. Optogenetic Suppression of D2-MSN in the NAc During Outcome Period**

865 **After Error Selectively Decreases Response Latency**

866 (A) Optogenetic suppression of D2-MSN in the NAc during ITI period did not affect response latencies after

867 correct or error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Treatment effects,

868 $F_{1,8} = 0.01032$, $p = 0.9216$; After Correct, $p = 0.1918$; After Error, $p = 0.1553$).

869 (B) Light stimulation in D2-eYFP during ITI period did not affect response latencies after correct or error trial

870 in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction, $F_{1,7} =$

871 $0.1309$, $p = 0.7281$; After Correct, $p = 0.9493$; After Error, $p = 0.9732$).

872 (C) Optogenetic suppression of D2-MSN in the NAc during Cue period did not affect response latencies after

873 correct or error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Treatment effects,

874 $F_{1,8} = 1.331$, $p = 0.2820$; After Correct, $p = 0.8452$; After Error, $p = 0.1203$).

875 (D) Light stimulation in D2-eYFP during Cue period did not affect response latencies after correct or error trial
876 in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction, $F_{1,7}$ =
877 0.0003527, p = 0.9855; After Correct, p = 0.9046; After Error, p = 0.9158).
878 (E) Optical stimulation of D2-MSN in the NAc during Outcome period after error response decreased response
879 latencies in the next trial of the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Treatment
880 effects, $F_{1,8}$ = 23.51, **p = 0.0013; After Correct, p = 0.9716; After Error, ***p = 0.0003).
881 (F) Light stimulation in D2-eYFP during Outcome period did not affect response latencies after correct or error
882 trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type × Treatment interaction,
883 $F_{1,7}$ = 0.4414, p = 0.5277; After Correct, p = 0.7563; After Error, p = 0.9668).
884
885 **Supplementary Figure S7. GRIN Lens Placement**
886 Histology on GRIN lens placement.
887
888 **Supplementary Figure S8. Histograms of auROC During ITI, Cue, and Outcome Period**
889 (A - C) Histogram of averaged auROC of D1-MSN during ITI (A), Cue (B), and Outcome (C) period.
890 (D - F) Histogram of averaged auROC of D2-MSN during ITI (D), Cue (E), and Outcome (F) period.
891
892 **Supplementary Figure S9. Averaged Calcium Traces from Responsive Type I and Type II Neurons of**
893 **D1-Cre and D2-Cre mice**
894 (A) Averaged traces of all responsive Type I neurons in D1-Cre mice (n = 35).
895 (B) Averaged traces of all responsive Type II neurons in D1-Cre mice (n = 38).
896 (C) Averaged traces of all responsive Type I neurons in D2-Cre mice (n = 43).
897 (D) Averaged traces of all responsive Type II neurons in D2-Cre mice (n = 15).
898 (E) A scatterplot of individual D1-MSN or D2-MSN responses during Outcome period.
899 (F) Averaged auROC traces of all responsive Type I neurons in D1-Cre and D2-Cre mice.
900 (G) Early responses (2.0 – 5.0 sec) of all responsive Type I (unpaired t-test, $t_{76}$ = 1.289, p = 0.2014).
901 (H) Late responses (7.0 – 10.0 sec) of all responsive Type I (unpaired t-test, $t_{76}$ = 2.373, *p = 0.0202).

902 (I) Averaged auROC traces of all responsive Type II neurons in D1-Cre and D2-Cre mice.

903 (J) Early responses (2.0 – 5.0 sec) of all responsive Type II (unpaired t-test, $t_{51} = 0.9313$, $p = 0.3561$).

904 (K) Late responses (7.0 – 10.0 sec) of all responsive Type II (unpaired t-test, $t_{51} = 0.1162$, $p = 0.9080$).

905

906 **Supplementary Figure S10. Neurons of the Same Type Stays Closer to Each Other than Neurons of**

907 **Different Type in D2-MSN**

908 (A and B) Spatial mapping of individual type neurons from representative D1-Cre (A) or D2-Cre (B) mice.

909 (C and D) Quantification of the pairwise distances of each types of neurons in D1-Cre (C, KS test, $p = 0.4503$)

910 or D2-Cre (D, KS test, $p = 0.5357$) mice.

911 (E and F) Quantification of the pairwise distances of different types of neurons in D1-Cre (E, KS test, $p =$

912 $0.2519$) or D2-Cre (F, KS test, ***$p = 0.0002$) mice.

913

914 **Supplementary Figure S11. The Behavioral Performance of Novice and Expert Mice**

915 Behavioral performances were improved through learning (Two-way RM-ANOVA with Sidak correction,

916 Learning effects, $F_{1,4} = 23.33$, **$p = 0.0085$; Novice, $p = 0.3465$; Expert, $p = 0.8630$).

917

918 **Supplementary Figure S12. Averaged auROC Traces of Each Type Neuron Through Learning**

919 (A) Averaged auROC traces of neurons becoming responsive Type I from non-responsive type neuron in D1-

920 Cre mice (Left). Averaged auROC in expert mice was significantly higher than that in novice mice (Right,

921 paired-test, $t_6 = 10.22$, ***$p < 0.0001$).

922 (B) Averaged auROC traces of neurons becoming responsive Type II from non-responsive type neuron in D1-

923 Cre mice (Left). Averaged auROC in expert mice was significantly lower than that in novice mice (Right,

924 paired-test, $t_{13} = 8.766$, ***$p < 0.0001$).

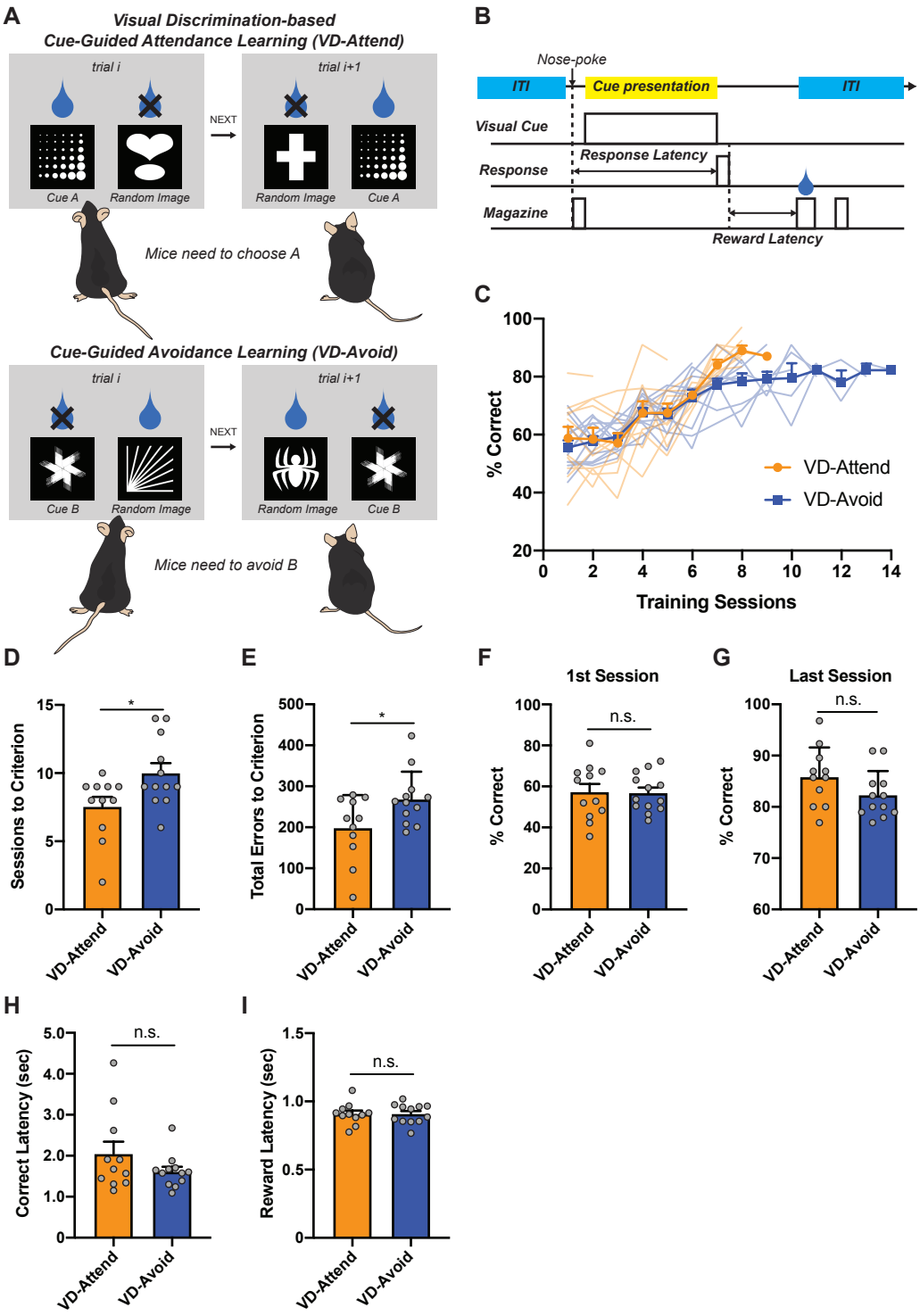
925 (C) Averaged auROC traces of neurons becoming responsive Type I from non-responsive type neuron in D2-

926 Cre mice (Left). Averaged auROC in expert mice was significantly higher than that in novice mice (Right,

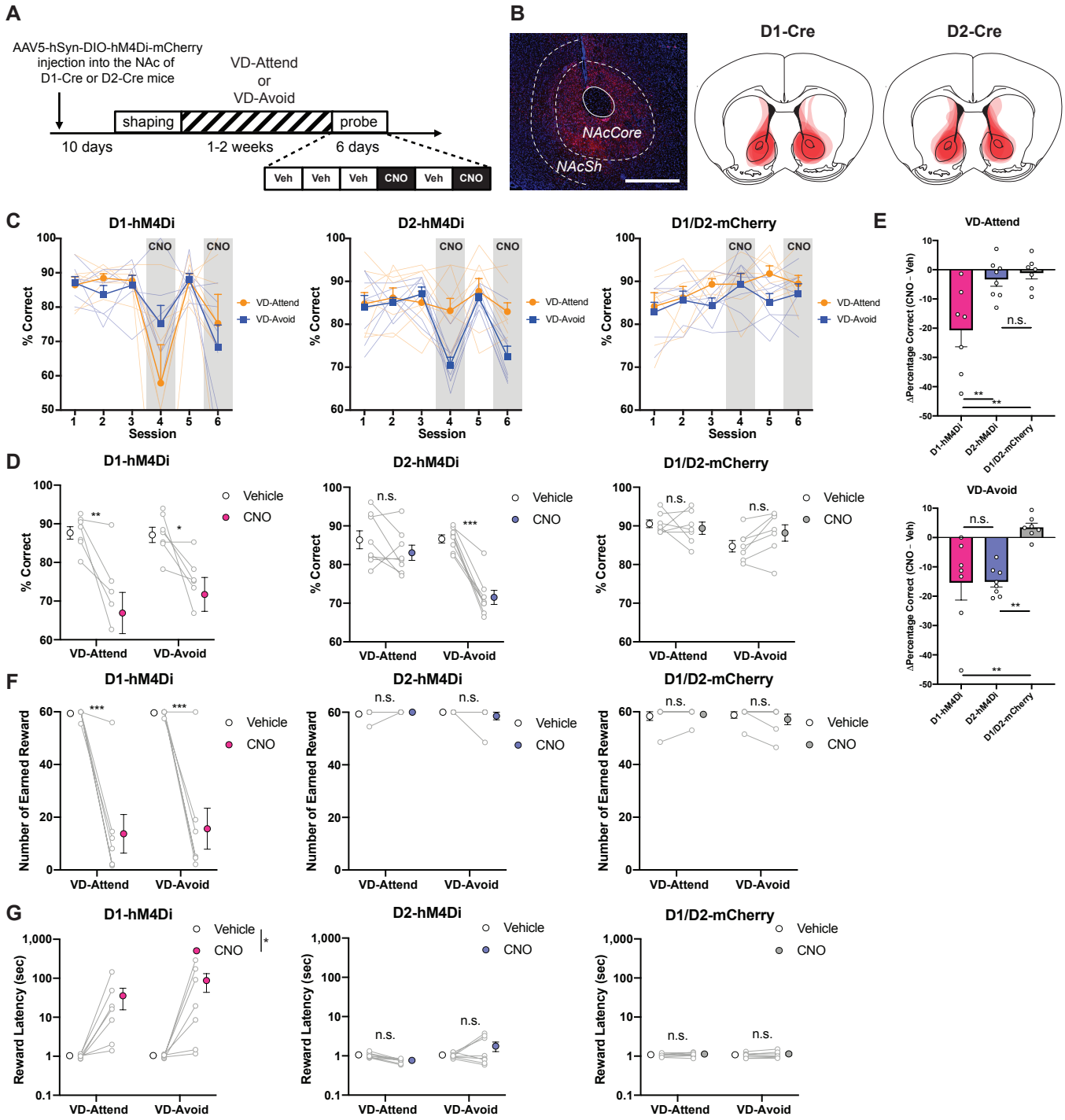
927 paired-test, $t_{21} = 7.782$, ***$p < 0.0001$).

928 (D) Averaged auROC traces of neurons becoming responsive Type II from non-responsive type neuron in D2-
929 Cre mice (Left). Averaged auROC in expert mice was significantly lower than that in novice mice (Right,
930 paired-test, $t_4 = 3.841$, *$p = 0.0184$).
931
932 **Supplementary Movie 1. Calcium Imaging from Mice Performing the VD-Avoid Task in Correct Trial**
933
934 **Supplementary Movie 2. Calcium Imaging from Mice Performing the VD-Avoid Task in Error Trial**
935
936

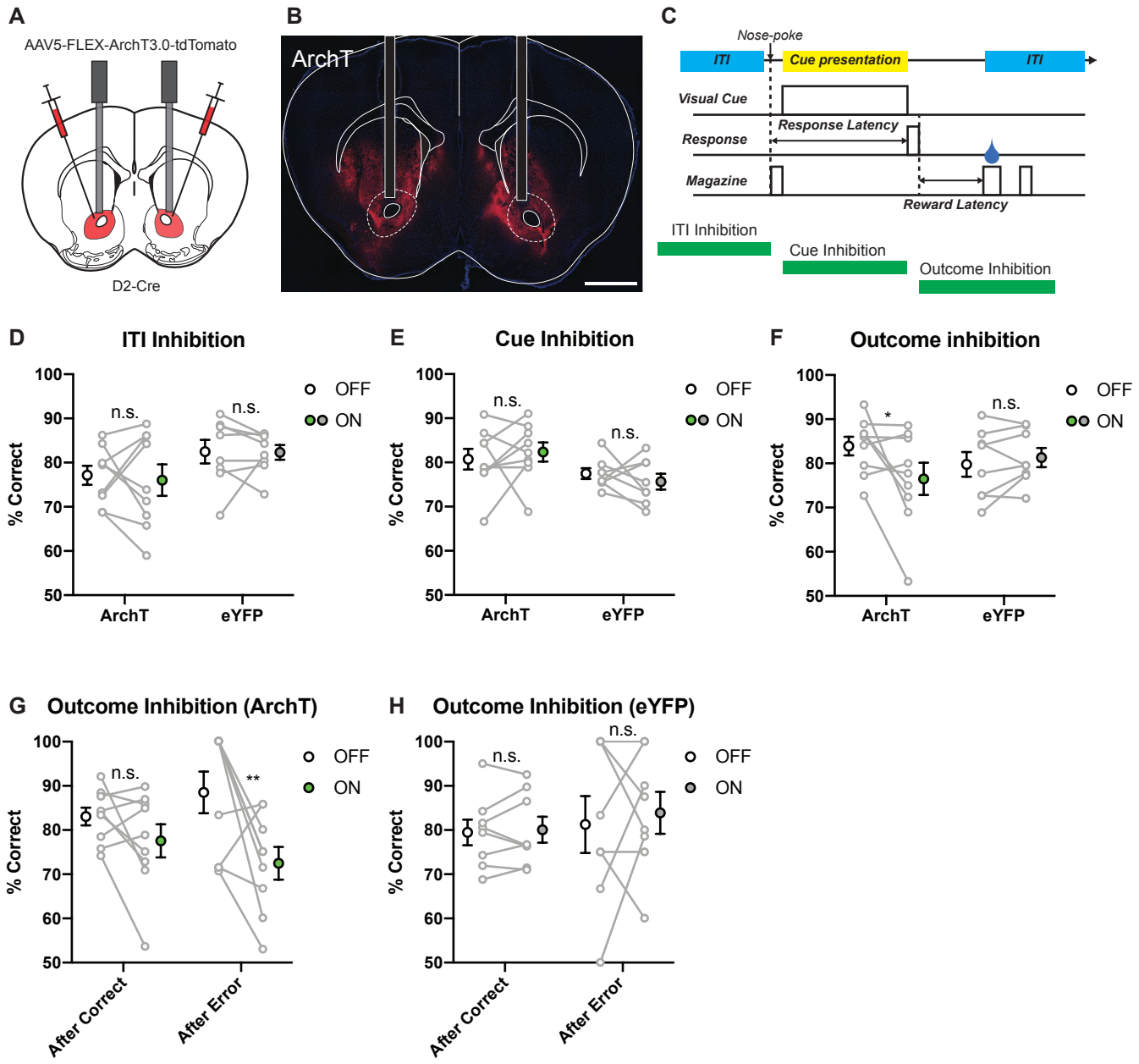
# Figure 1 Nishioka et al.



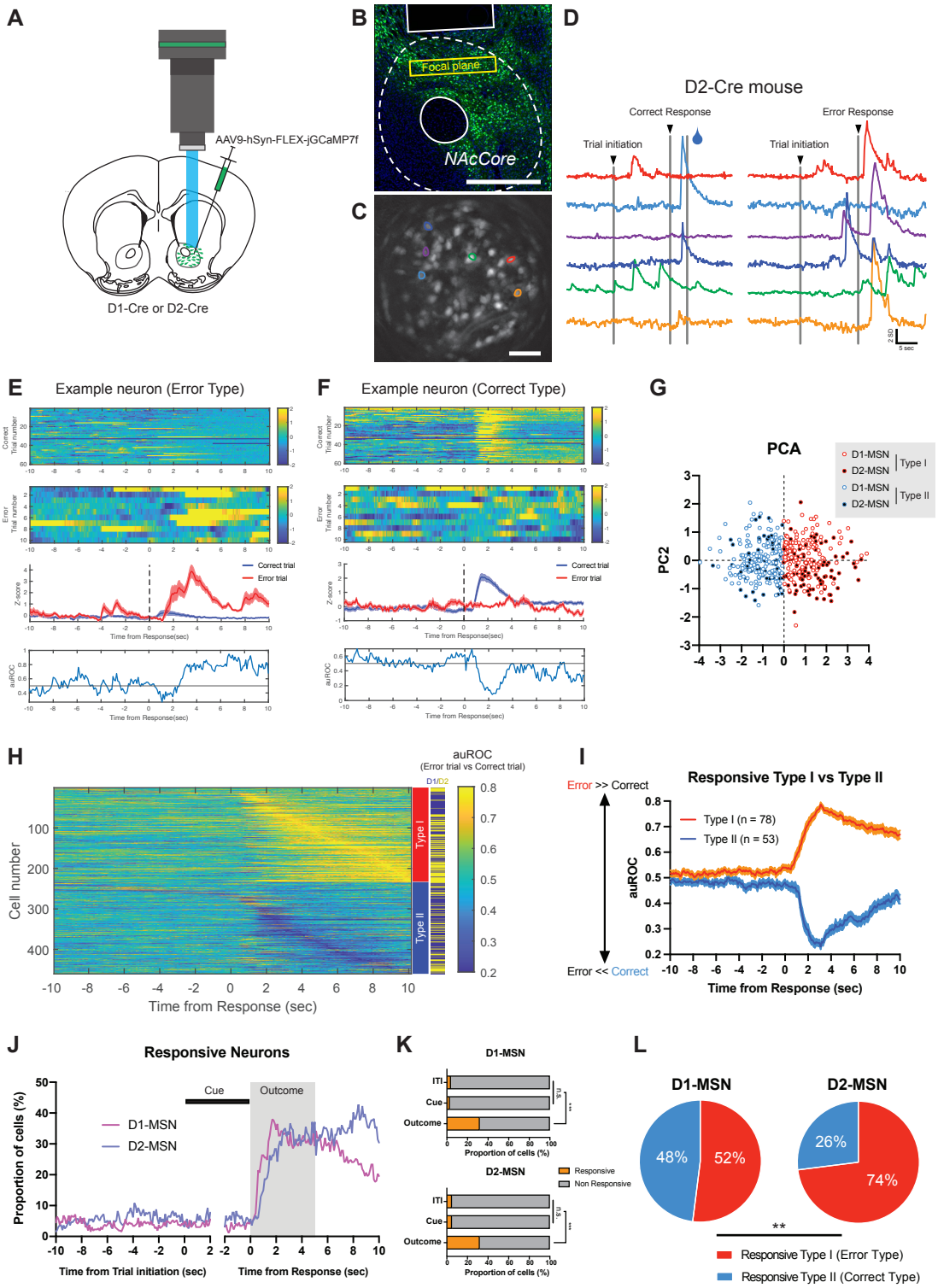
# Figure 2 Nishioka et al.



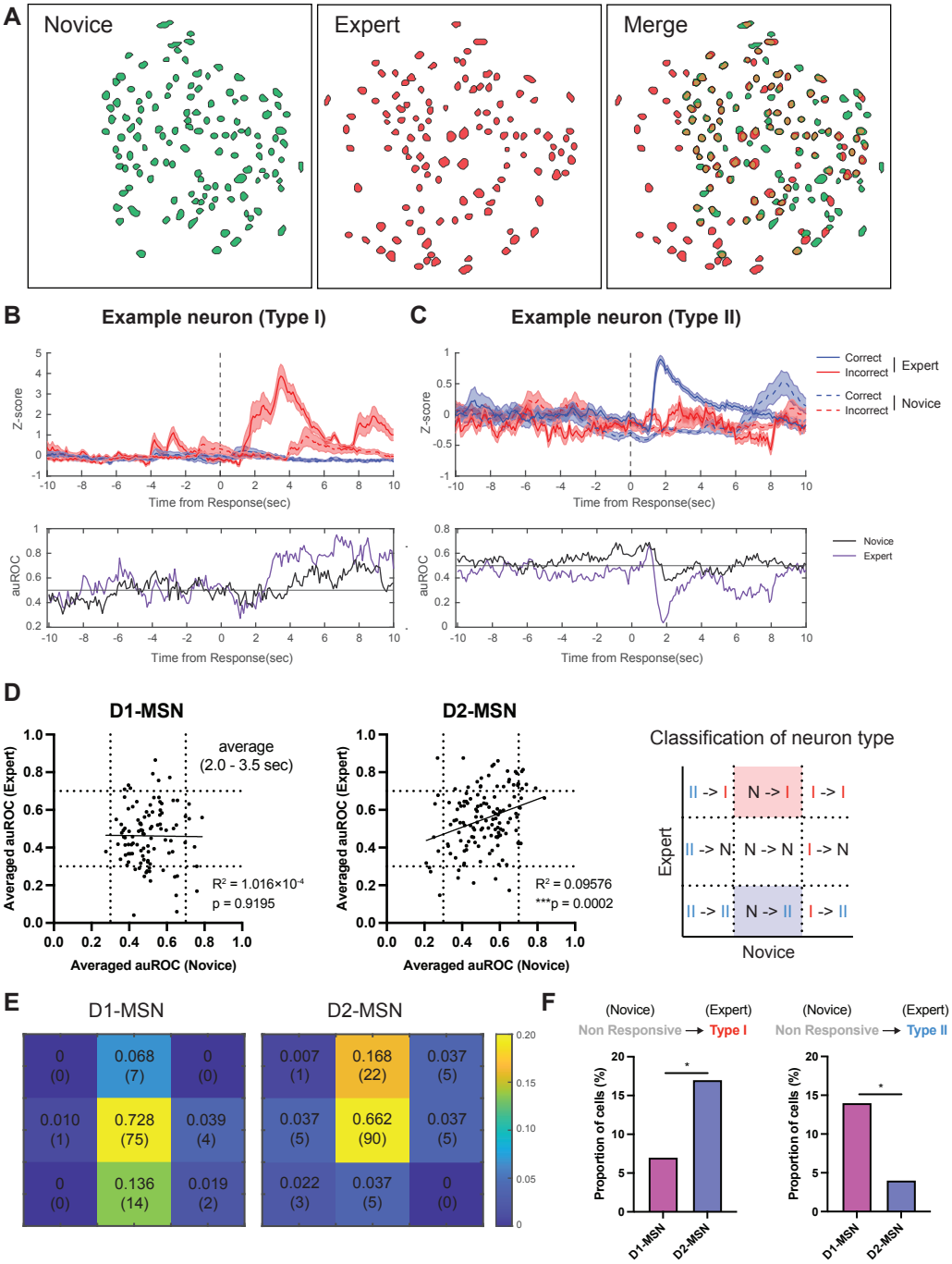
# Figure 3 Nishioka et al.



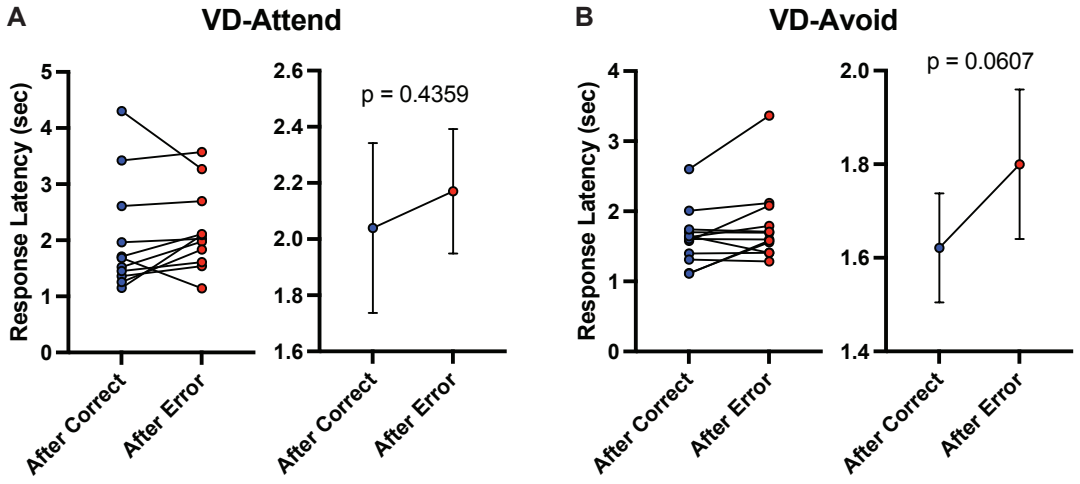
# Figure 4 Nishioka et al.



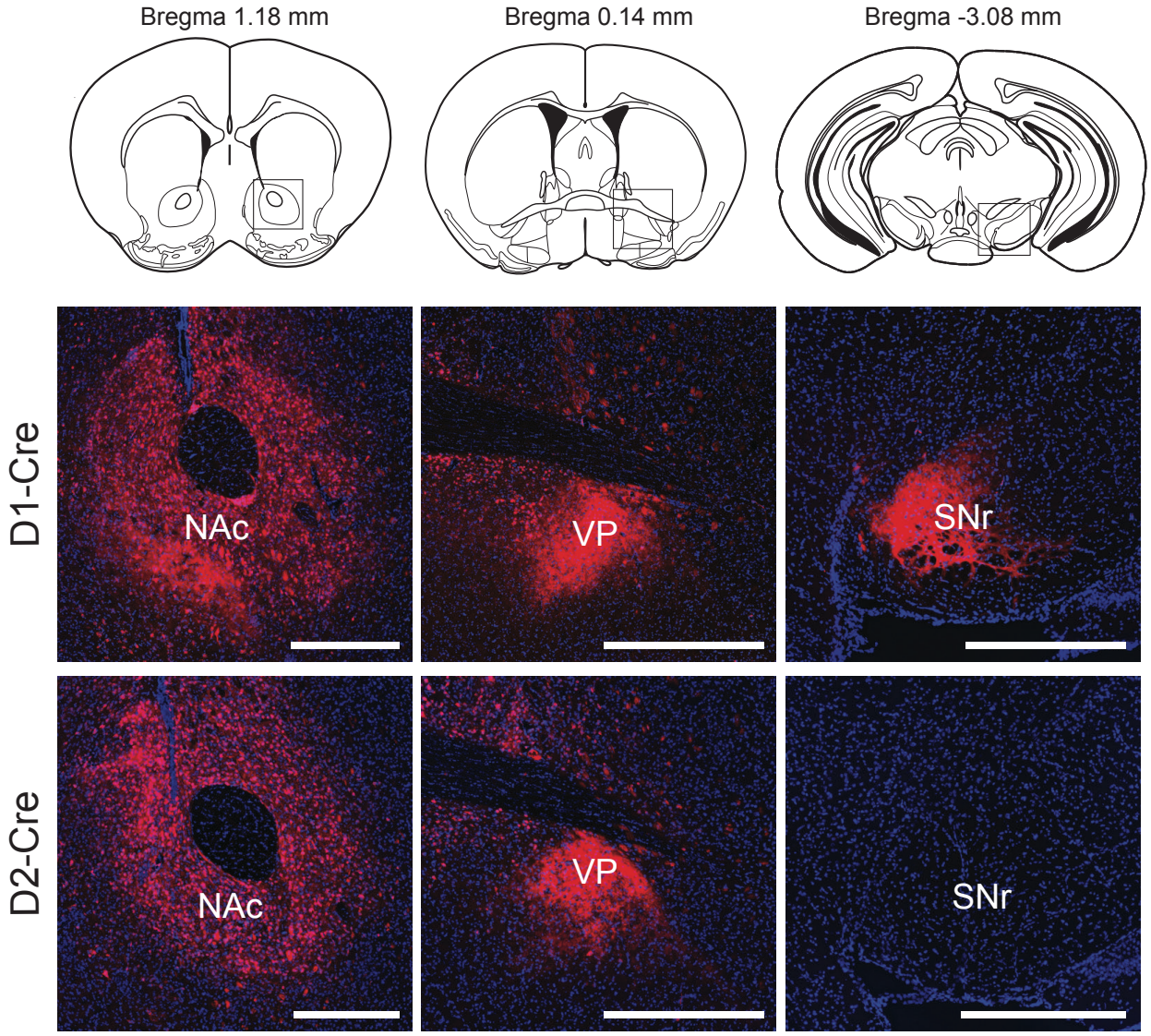
# Figure 5 Nishioka et al.



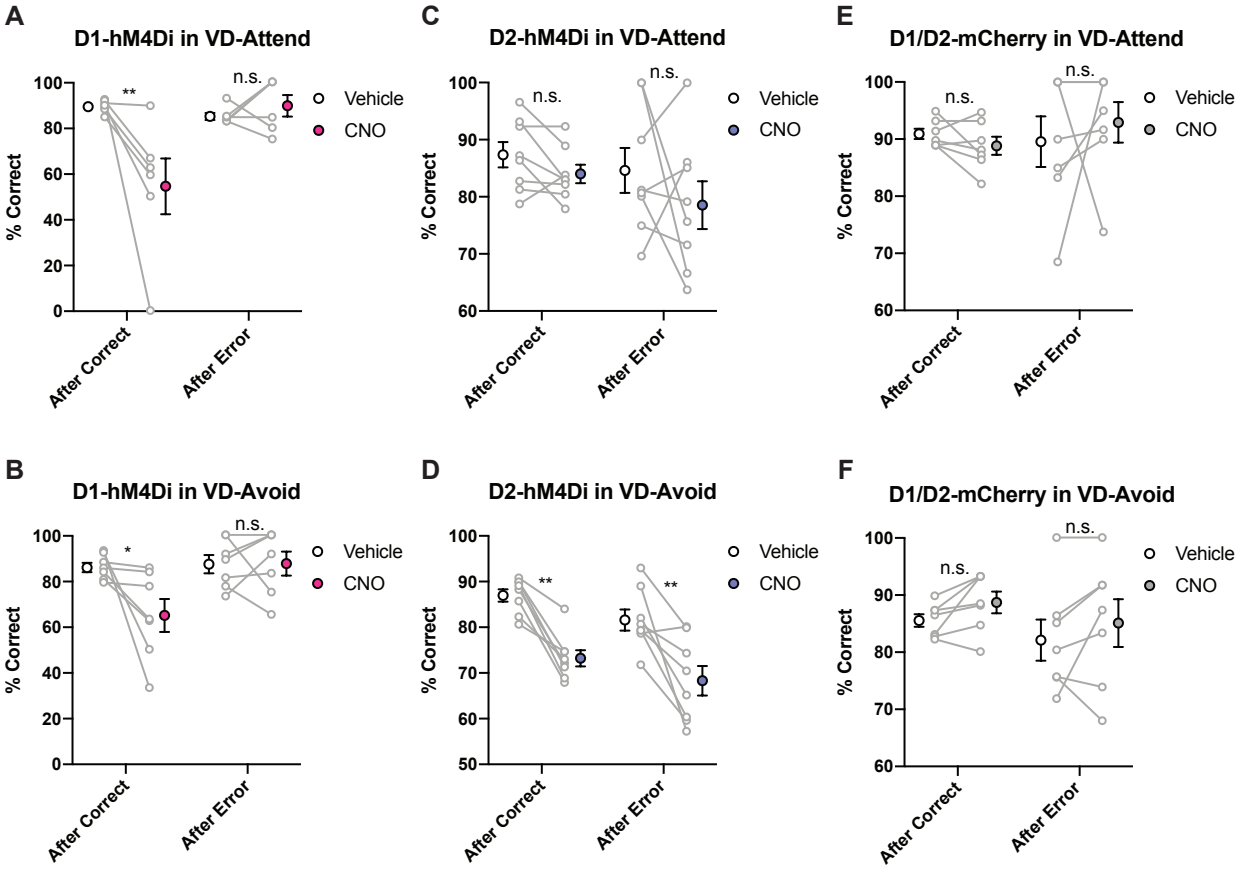
Supplementary Figure S1 Nishioka et al.



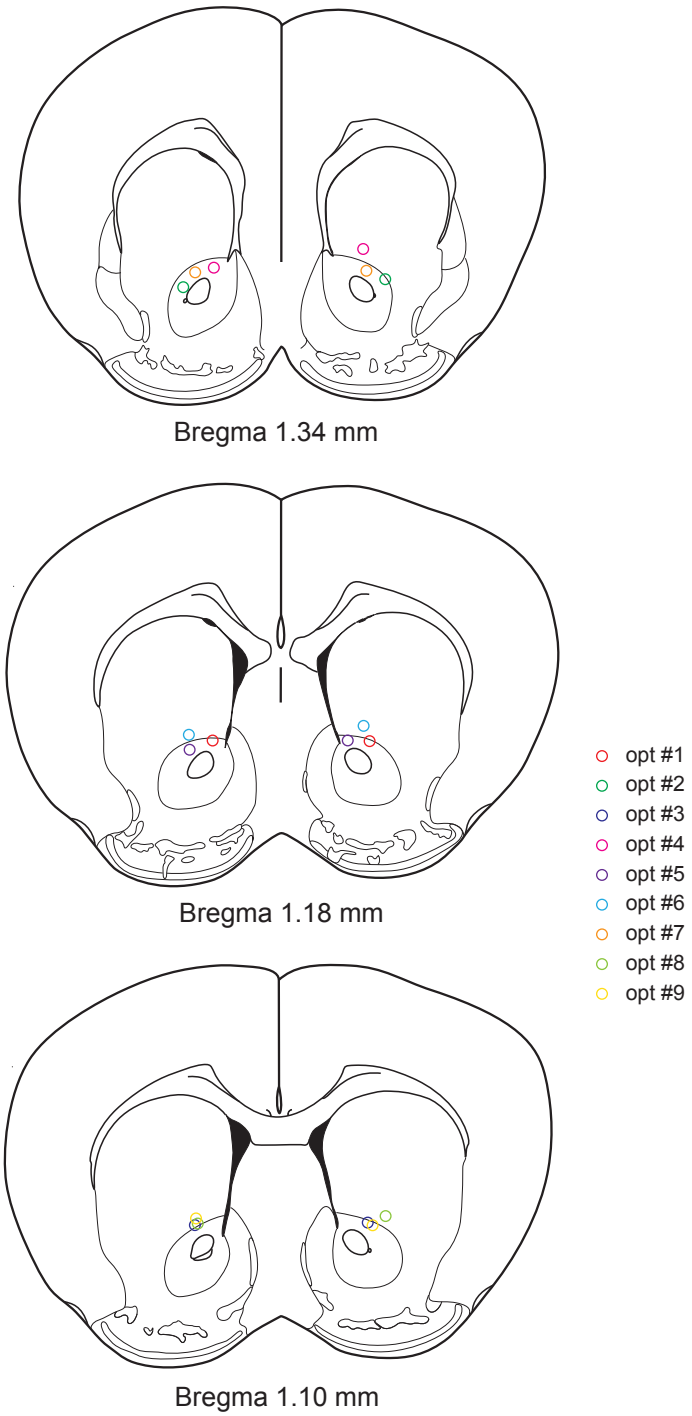
Supplementary Figure S2 Nishioka et al.



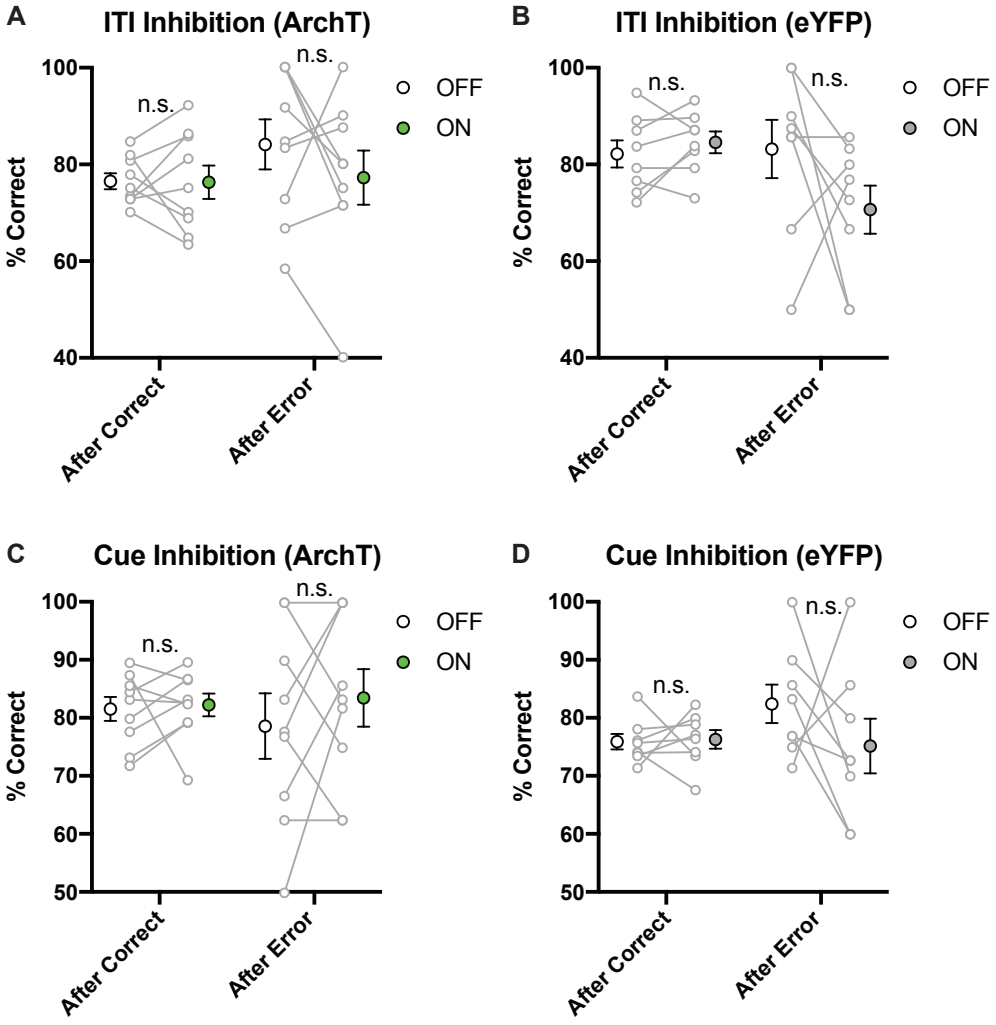
Supplementary Figure S3 Nishioka et al.



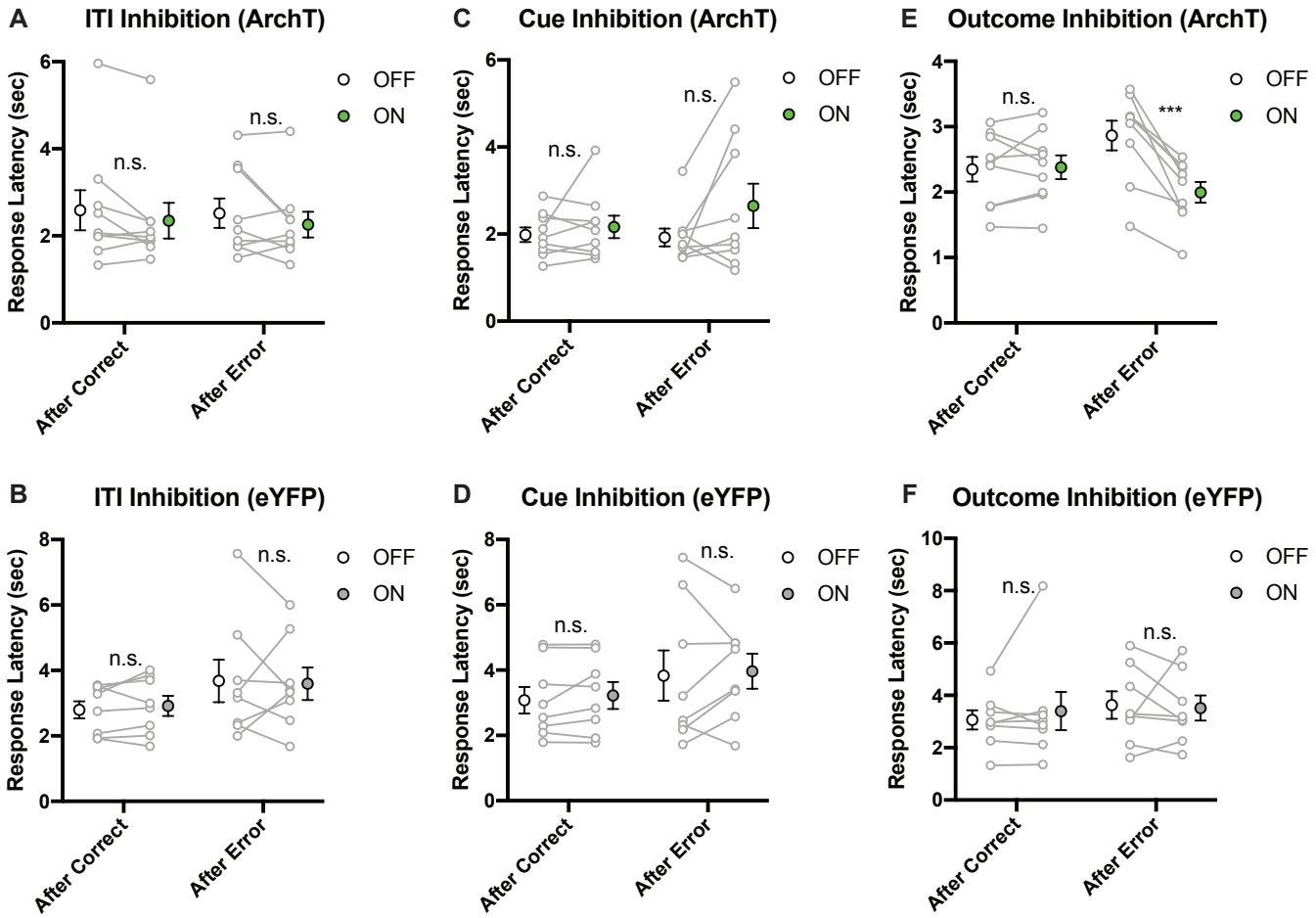
# Supplementary Figure S4 Nishioka et al.



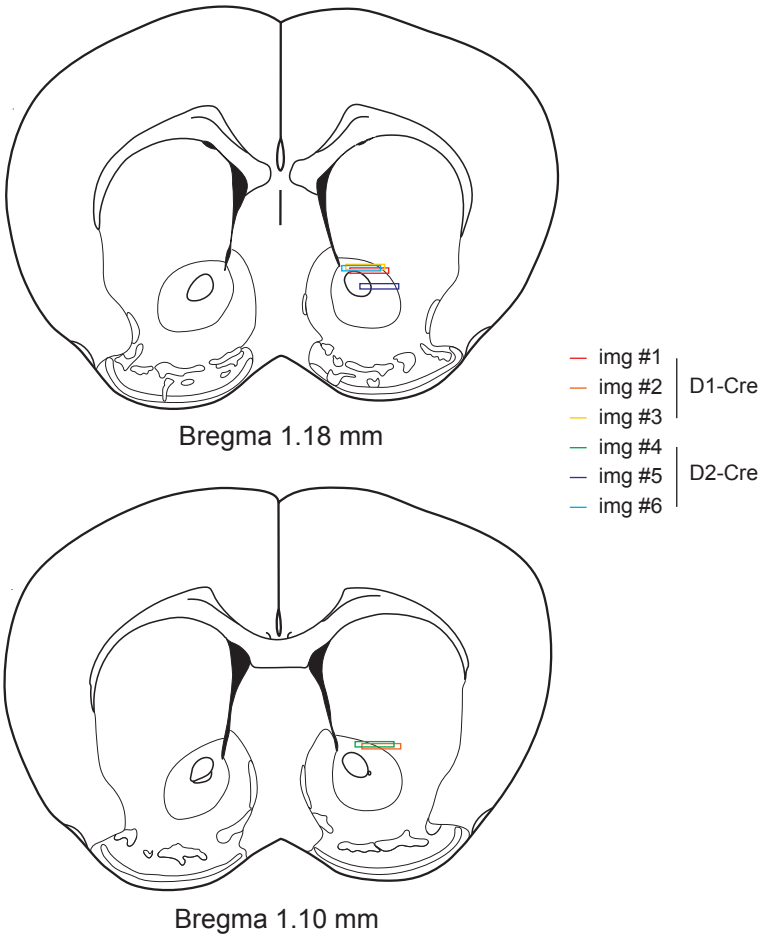
Supplementary Figure S5 Nishioka et al.



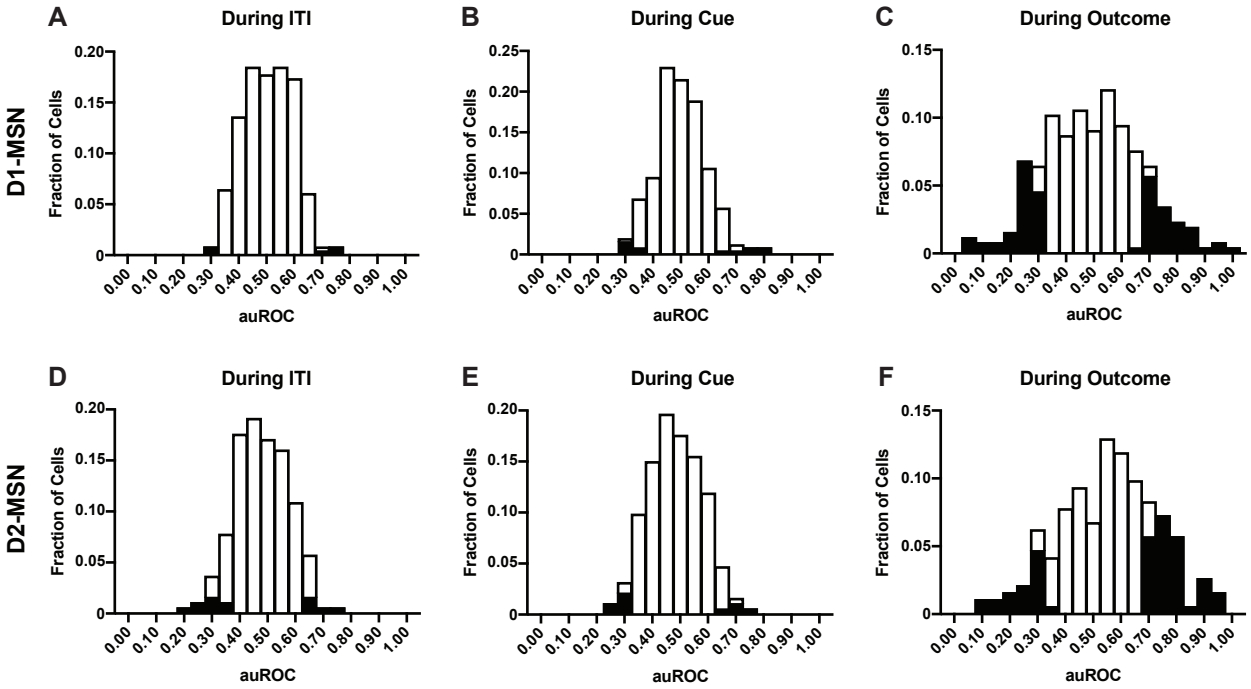
# Supplementary Figure S6 Nishioka et al.



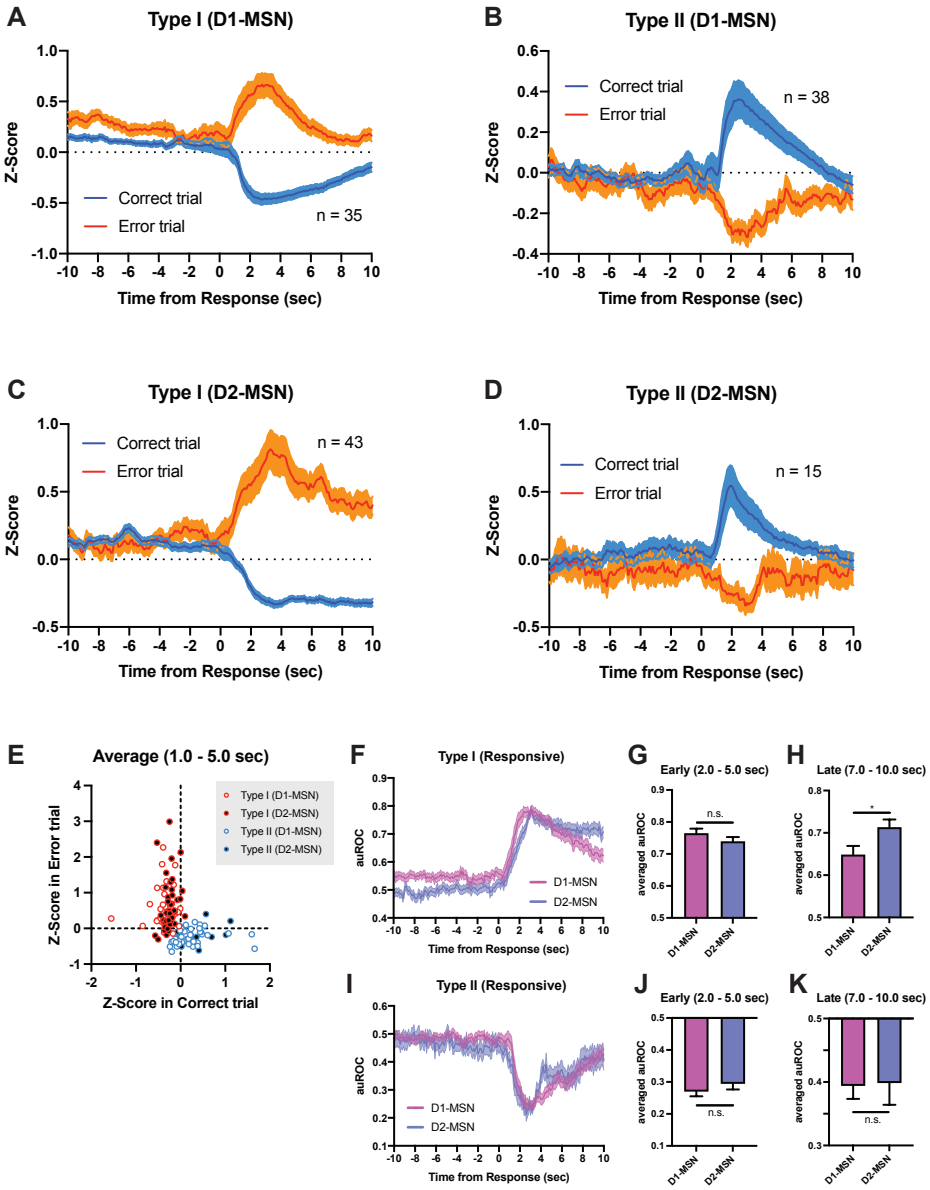
# Supplementary Figure S7 Nishioka et al.



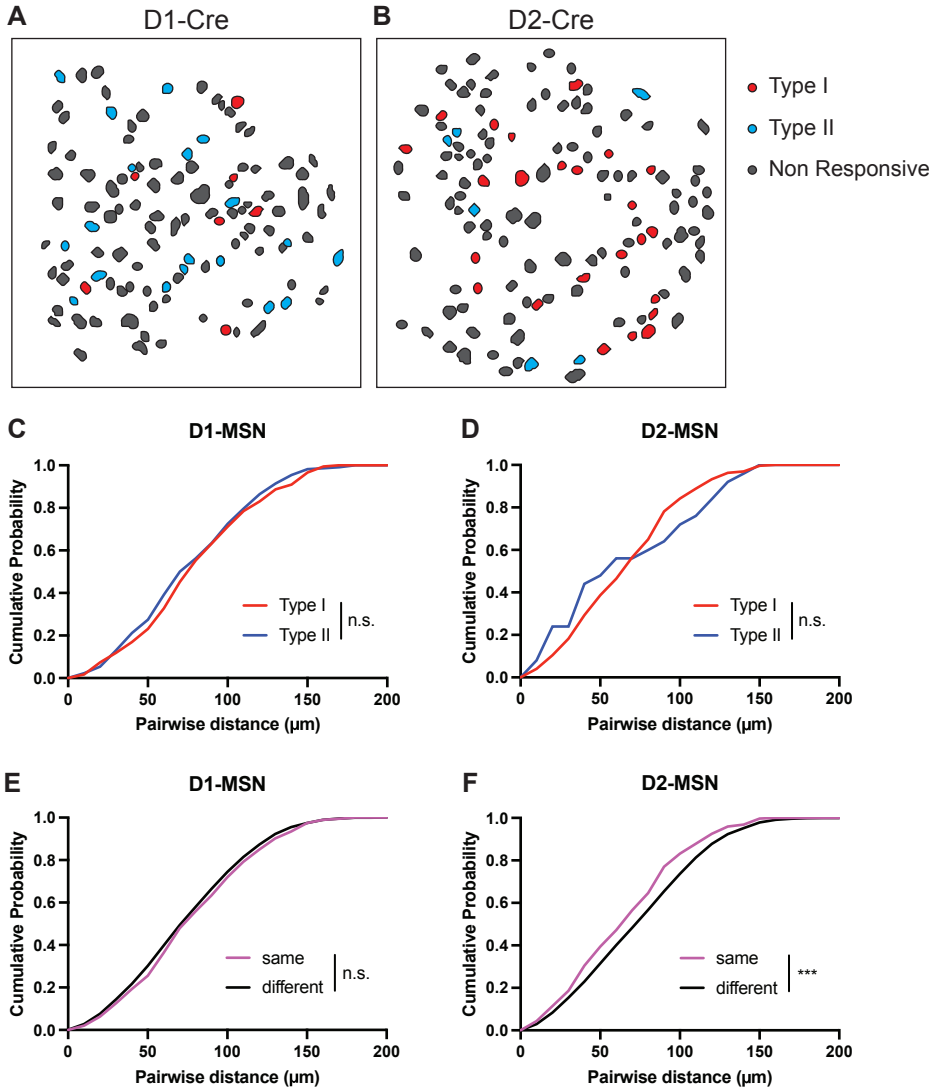
Supplementary Figure S8 Nishioka et al.



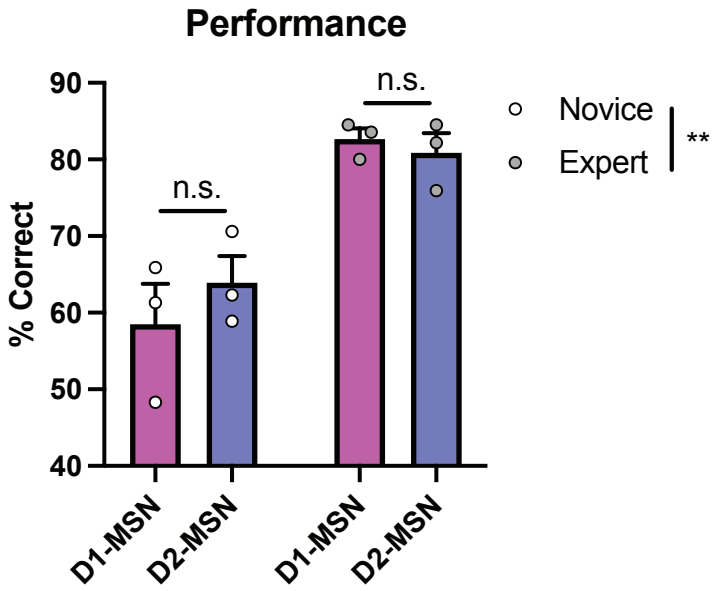
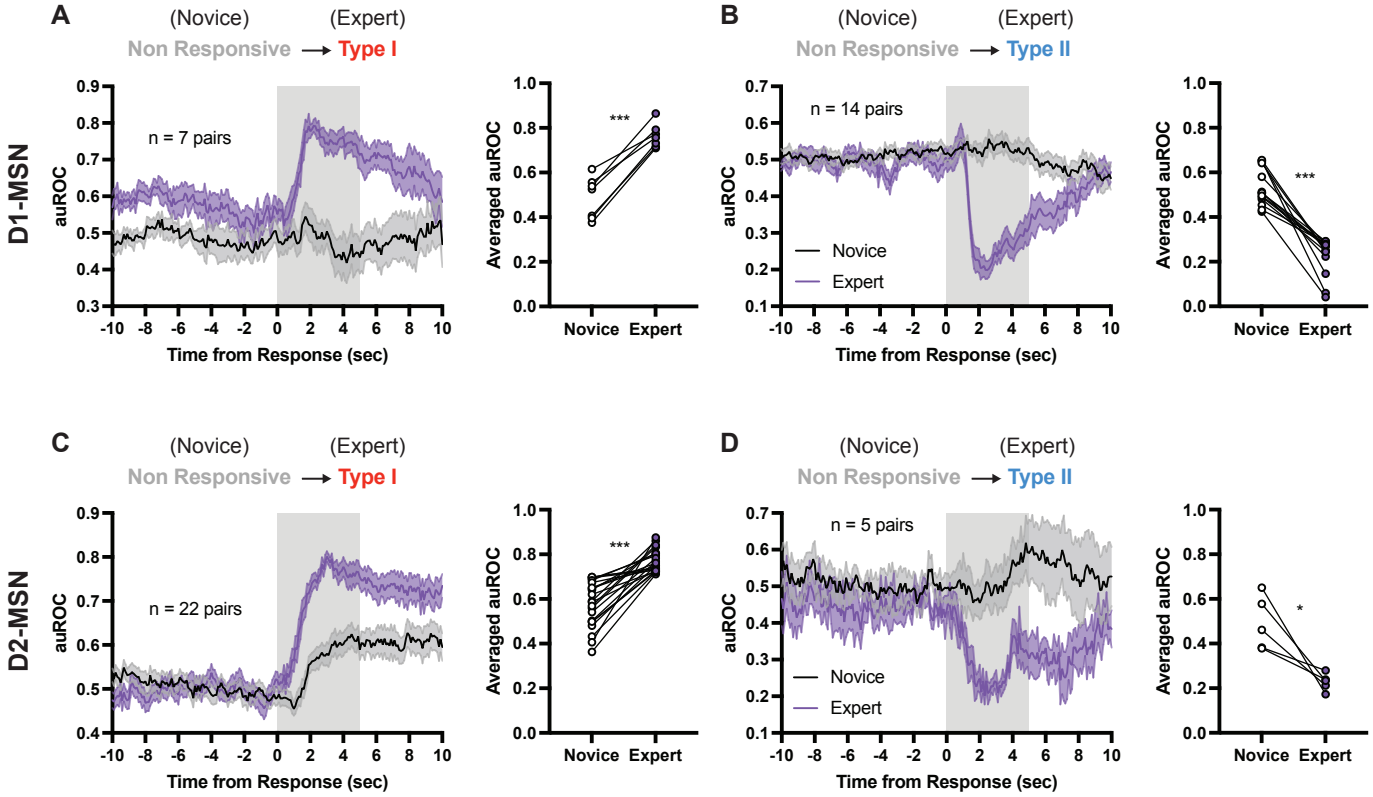
# Supplementary Figure S9 Nishioka et al.



# Supplementary Figure S10 Nishioka et al.



Supplementary Figure S11 Nishioka et al.



# Supplementary Figure S12 Nishioka et al.